

Critical exponents and fastest convergence rates of distributed consensus with switching topologies and additive noises*

Ge Chen,^{*} Le Yi Wang,[†] Chen Chen,[‡] George Yin[§]

Abstract

Conditions for convergence and convergence speed are crucial for distributed consensus algorithms of networked systems. Based on a basic first-order average-consensus protocol, this paper estimates its fastest convergence rate under unknown switching network topologies and additive noises. It is shown that the fastest convergence rate is of the order $1/t$ for the best topologies, and for the worst topologies which are balanced and uniformly jointly connected. This implies for any balanced and uniformly jointly connected topologies, the fastest convergence rate is of the order $1/t$. Another effort of this paper is to investigate the critical topological condition for consensus of this protocol. We propose a new joint-connectivity condition called extensible joint-connectivity that contains a parameter δ (termed as an extensible exponent). Using this condition we show the critical value of δ for consensus is $1/2$, and prove the fastest convergence speed is of the order $t^{1-2\delta}$, which quantitatively describes a relation between the convergence speed and connectivity. Also, the extensible joint-connectivity can be used to deal with the system whose network topology affects complex uncertainty, such as the consensus analysis of the system with non stationary and strongly correlated stochastic topologies.

Keywords: average-consensus, stochastic approximation, jointly-connected topology, multi-agent system, networked system

*The research of G. Chen and C. Chen was supported by the National Key Basic Research Program of China (973 program) under grant 2014CB845301/2/3, and by the National Natural Science Foundation of China under grants 91427304. The research of L.Y. Wang and G. Yin was supported in part by U.S. Air Force Office of Scientific Research under FA9550-15-1-0131.

^{*}National Center for Mathematics and Interdisciplinary Sciences & Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China, chenge@amss.ac.cn

[†]Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202, USA, lywang@ece.eng.wayne.edu

[‡]Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China, chenchen@amss.ac.cn

[§]Department of Mathematics, Wayne State University, Detroit, MI 48202, USA, gyin@math.wayne.edu

1 Introduction

Consensus of multi-agent systems has drawn considerable attention from various fields over the past two decades. For example, physicists investigate the synchronization phenomena of coupled oscillators, flashing fireflies, and chirping crickets [1–3]; biologists, physicists, and computer scientists try to understand and model the flocking phenomenon of animals behavior [4–6]; sociologists simulate the emergence and spread of the public opinions [7, 8]. Because of the importance, effort has been devoted to the mathematical analysis of consensus of flocks [9–12]. Meanwhile consensus control algorithms have been developed for a wide range of applications such as the formation controls of robots and vehicles [13–16], the attitude synchronization of rigid bodies and multiple spacecraft [17–19], and the distributed computation, filtering and resource allocations of networked systems [20, 21, 41], etc. The common thread in the consensus research is a group of agents with interconnecting neighbor graphs try to achieve a global coordination or collective behavior by using neighborhood information permitted by the network topologies.

The first-order average-consensus algorithm is the most basic consensus protocol. Such an algorithm is investigated by many different manners because the actual systems are affected by different kinds of uncertainties. For example, in the wireless communication networks the communication is affected by the thermal noise, channel fading and quantization effect; in the formations of multiple satellites, vehicles or robots, there exist measurement noises in the observation of the neighbors' states. To handle the random failure of the communication links, some papers use the deterministic network topology but allowing it switching [22–24], and some others adopt the stochastic setting where the network topology evolves according to some random distributions [25–30]. We remark that the models in these research contain no noise item, which means these research cannot handle some uncertainties such as the measurement error and quantization noise. To overcome this default some papers consider the first-order average-consensus protocols with additive noises [31–44], among which a typical method is to use a stochastic approximation methods.

The main idea of the method of distributed stochastic approximation is: Each agent in the network uses a decreasing gain function acting on the information received from its neighbors to reduce the communication or measurement noises. Using this idea, Huang and Manton [36] and Li and Zhang [39] considered the first-order discrete-time and continuum-time consensus model with fixed topology and additive noises respectively. They showed that the system can achieve consensus in probability sense if the topology is balanced and connected. Later, this consensus condition had been relaxed from fixed balanced topology to switched balanced topologies satisfying a *uniform joint-connectivity* condition [31, 40], i.e., the union of the topology graphs over every bounded time interval is connected, and further relaxed to the general directed topologies with uniform joint-connectivity [38]. Huang [37] also applied the stochastic approximation to the consensus problems for over lossy wireless network containing random link gains, additive noises and Markovian lossy signal receptions. On the other hand, motivated by resource allocation problems in computing, communications, inventory, space, and power generations, Yin, Sun and Wang [41] introduced a stochastic approximation algorithm for constrained consensus problem of networked system, where the consensus conditions are established by assuming the topologies are randomly switched under a Markov chain framework. This algorithm was further investigated and expanded by some literatures later [42–44].

Despite the existing research work on the first-order average-consensus algorithms, some key problems remain unsolved. First, from [30] it is shown that the first-order average-

consensus protocol with deterministic topologies and no additive noises will achieve consensus if and only if the time-varying topologies satisfying an *infinite joint-connectivity* condition, i.e. the union of the topologies from any finite time to infinite is connected, providing all the topologies have the same stationary distribution. However, for the same protocol but with additive noises the current best condition of topologies for consensus is the uniform joint-connectivity [31, 38, 40]. Thus, there exists a huge gap between the consensus condition of topologies with and without additive noises. Naturally, for the protocol with additive noises, an open question is: What is the critical condition of the topologies for consensus? Second, convergence speed is an important performance of the consensus protocols. For the noise-free algorithms it has been well investigated [22–25, 27, 28, 45, 46] while some literature tries to maximize the convergence speed by optimizing the weighted network topology [23, 47]. However, few work considers the convergence speed of the systems containing uncontrollable time-varying network topologies and additive noises. Furthermore, to our best knowledge no existing paper investigates how fast the fastest convergence rate of such kind systems can achieve by optimizing their own controllers.

This paper addresses the above problems for the first-order average-consensus algorithm with time-varying network topologies and additive noises. Our algorithm assumes that each node only knows its own and neighbors' information and the network topologies cannot be real-time controlled. We use the stochastic approximation methods to attenuate the affection of the noises and treat the gain function as the only control input. The main contribution of this paper can be listed as follows:

- We propose a new condition for topologies named *extensible joint-connectivity* (to be precisely defined later in the paper). Our conditions allows the length of the interval during which the union of the network topologies is connected can increase to infinity as time grows to infinity, which is an intermediate condition between the uniform joint-connectivity and infinite joint-connectivity. The extensible joint-connectivity contains a key parameter named *extensible exponent* denoted by δ whose value corresponds to the rate of increments of the length of the interval guaranteeing joint-connectivity. Under this new condition, we show that the critical value of δ for designing the gain function to guarantee consensus is $\frac{1}{2}$. We also derive a critical value about the expectations of the noises. As an application, our results are used to analyze the systems with complex random network topologies like non stationary and strongly correlated random sequence.
- We first investigate the fastest speed to reach consensus of our systems among all the gain functions that is called as the *fastest convergence rate*. Note that the fastest convergence rate depends on the time-varying network topologies and the noises. Under some noise conditions, we show the fastest convergence rate is of the order $(1/t)$ (i.e., $\Theta(\frac{1}{t})$) for the best topologies, and also for the worst topologies which are balanced and uniformly jointly connected. This indicates for any balanced and uniformly jointly connected topologies the fastest convergence rate is $\Theta(\frac{1}{t})$. Moreover, if the network topologies satisfy the extensible joint-connectivity condition, it is shown that the fastest convergence rate is $\Theta(t^{1-2\delta})$, which quantitatively describes a relation between the convergence speed and connectivity. We remark that these investigations differ from previous works [23, 47], which assume the network topology is controllable and optimize it to maximize the convergence rate.

The rest of the paper is organized as follows: Section 2 introduces the consensus protocol and some basic definitions. Section 3 investigates the consensus conditions including the sufficient conditions and necessary conditions, and Section 4 provides the convergence speed

including the lower bounds and upper bounds. According these two sections we sum up our main results in Section 5. Finally, we conclude this paper in the last section.

2 Preliminaries

2.1 Definitions in Graph Theory

Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathcal{A}\}$ be a weighted digraph, where $\mathcal{V} = \{1, 2, \dots, n\}$ is the set of nodes with node i representing the i th agent, \mathcal{E} is the set of edges, and $\mathcal{A} \in \mathbb{R}^{n \times n}$ is the weight matrix. An edge in \mathcal{G} is denoted by an ordered pair (j, i) , and $(j, i) \in \mathcal{E}$ if and only if the j th agent can send information to the i th agent directly. The neighborhood of the i th agent is denoted by $\mathcal{N}_i = \{j \in \mathcal{V} | (j, i) \in \mathcal{E}\}$. An element of \mathcal{N}_i is called a neighbor of i . Let $a_{ij} > 0$ denote the weight of the information from node j to i . The weighted matrix \mathcal{A} is defined by $\mathcal{A}_{ij} = a_{ij}$ for $1 \leq i \neq j \leq n$, and $\mathcal{A}_{ii} = 0$ for $1 \leq i \leq n$.

For graph \mathcal{G} , the in-degree of i is defined as $\deg_{\text{in}}^i = \sum_{j \in \mathcal{N}_i} a_{ij}$ and the out-degree of i is defined as $\deg_{\text{out}}^i = \sum_{j \in \mathcal{N}_i} a_{ji}$. If $\deg_{\text{in}}^i = \deg_{\text{out}}^i$ for all $1 \leq i \leq n$, we call \mathcal{G} a balanced digraph. The Laplacian matrix of \mathcal{G} is defined as $L_{\mathcal{G}} = D_{\mathcal{G}} - \mathcal{A}_{\mathcal{G}}$, where $D_{\mathcal{G}} = \text{diag}(\deg_{\text{in}}^1, \dots, \deg_{\text{in}}^n)$, and $[\mathcal{A}_{\mathcal{G}}]_{ij}$ equals to a_{ij} if $j \in \mathcal{N}_i$ and 0 otherwise. \mathcal{G} is called an undirected graph, if $\mathcal{A}_{\mathcal{G}}$ is a symmetric matrix. It is easily shown that an undirected graph must be a balanced digraph.

A sequence $(i_1, i_2), (i_2, i_3), \dots, (i_{k-1}, i_k)$ of edges is called a directed path from node i_1 to node i_k . The \mathcal{G} is called a strongly connected digraph, if for any $i, j \in \mathcal{V}$, there is a directed path from i to j . A strongly connected undirected graph is also called a connected graph. For graphs $\mathcal{G}(t) = \{V, \mathcal{E}(t), \mathcal{A}\}$, $i \leq t < j$, their union is defined by $\cup_{i \leq t < j} \mathcal{G}(t) := \{V, \cup_{i \leq t < j} \mathcal{E}(t), \mathcal{A}\}$. Note that there may exist multi edges from one vertex to another in $\cup_{i \leq t < j} \mathcal{G}(t)$.

2.2 Consensus Protocol

This paper considers a discrete-time first-order system containing n agents, where each agent i 's state $x_i(t)$ is updated by

$$x_i(t+1) = x_i(t) + u_i(t), \quad t = 1, 2, \dots \quad (2.1)$$

Here $u_i(t)$ is the control input of the i th agent. For simplicity, we suppose $x_i(t)$ and $u_i(t)$ are scalars. As mentioned above, this paper will investigate a basic stochastic approximation algorithm for average-consensus, that is, the control $u_i(t)$ is chosen by

$$u_i(t) = a(t) \sum_{j \in \mathcal{N}_i(t)} a_{ij} [x_j(t) - x_i(t) + w_{ji}(t)], \quad (2.2)$$

where $a(t) \geq 0$ is the common gain control at time t , $\mathcal{N}_i(t)$ is the neighbors of node i at time t , a_{ij} is the fixed weight from nodes j to i , $w_{ji}(t)$ is the noise of agent i receiving information from agent j at time t . Without loss of generality we assume $1 \leq a_{ij} \leq a_{\max}$, $1 \leq i \neq j \leq n$, where a_{\max} is a constant denoting the upper bound of a_{ij} . This kind protocol had been investigated by several papers [31, 36, 38–40] and has some applications in distributed computation of wireless sensor networks and formation controls of multiple satellites, vehicles or robots.

We define the σ -algebra generated by the noises $w_{ji}(k)$, $1 \leq k \leq t$, $1 \leq i \leq n$, $j \in \mathcal{N}_i(k)$ by

$$\mathcal{F}_t = \sigma(w_{ji}(k), 1 \leq k \leq t, 1 \leq i \leq n, j \in \mathcal{N}_i(k)).$$

The probability space of the system (2.1)-(2.2) is written to be $(\Omega, \mathcal{F}_\infty, P)$. Also, we write $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t), \mathcal{A})$ to be the topology of system (2.1)-(2.2) at time t , where $\mathcal{E}(t) = \{(j, i) | j \in \mathcal{N}_i(t)\}$ is the edge set of $\mathcal{G}(t)$. The corresponding topology sequence $\{\mathcal{G}(t)\}_{t \geq 1} = \{(\mathcal{V}, \mathcal{E}(t), \mathcal{A})\}_{t \geq 1}$. To simplify the exposition, we take $L(t) = L_{\mathcal{G}(t)}$ to be the Laplacian matrix of $\mathcal{G}(t)$. Let $x(t) = (x_1(t), x_2(t), \dots, x_n(t))'$. The protocol (2.1)-(2.2) can be rewritten as the following matrix form:

$$x(t+1) = [I - a(t)L(t)]x(t) + a(t)\hat{w}(t), \quad t = 1, 2, \dots, \quad (2.3)$$

where $\hat{w}(t) \in \mathbb{R}^n$ whose i -th element is $\sum_{j \in \mathcal{N}_i(t)} a_{ij}w_{ji}(t)$.

In this paper we assume there does not exist a central controller who knows the global information during the system's evolution, and the so-called consensus control means to design off-line gains $a(t)$ such that all agents achieve an agreement on their state in mean square sense, when $t \rightarrow \infty$. The mean square consensus is defined as follows:

Definition 2.1 *We say the system (2.1)-(2.2) reaches consensus in mean square if there exists a random variable x^* satisfying $|E(x^*)| < \infty$ and $\text{Var}(x^*) < \infty$ such that*

$$\lim_{t \rightarrow \infty} E\|x(t) - x^* \mathbf{1}\|^2 = 0,$$

where $\mathbf{1} \in \mathbb{R}^n$ is the column vector of all 1s.

2.3 Standard notation

The following standard mathematical notation will be used in this paper. Given a random variable X , let $E[X]$ and $\text{Var}(X)$ be its expectation and variance, respectively. For a vector Y , let Y_i denote its i th entry. For a real number x , $\lfloor x \rfloor$ is the maximum integer less than or equal to x , and $\lceil x \rceil$ is the smallest integer larger than or equal to x . Let $\|\cdot\|$ denotes the l_2 -norm (Euclidean norm). Given two sequences of positive numbers $g_1(t)$, $g_2(t)$,

- $g_1(t) = O(g_2(t))$ if there exists a constant $c > 0$ and a value $t_0 > 0$ such that $g_1(t) \leq cg_2(t)$ for all $t \geq t_0$.
- $g_1(t) = \Theta(g_2(t))$ if there exist two constants $c_2 > c_1 > 0$ and a value $t_0 > 0$ such that $c_1 g_2(t) \leq g_1(t) \leq c_2 g_2(t)$ for all $t \geq t_0$.
- $g_1(t) = o(g_2(t))$ if $\lim_{t \rightarrow \infty} \frac{g_1(t)}{g_2(t)} = 0$.

3 Consensus Conditions

This section will provide some consensus conditions for the system (2.1)-(2.2). In Subsection 3.1 we will propose a new condition concerning the connectivity, while the sufficient conditions and necessary conditions of consensus are put in Subsections 3.2 and 3.3 respectively.

3.1 Extensible Joint-connectivity

The uniform joint-connectivity of the topologies is a widely used important condition in the consensus research of multi-agent systems. However, such condition is not robust for some situations. For example, if the links in the network of a system have a positive probability of failure, it can be computed that with probability 1 the uniform joint-connectivity condition is not satisfied. Also, this condition cannot be satisfied in some flocking models [48]. To address more actual scenes we propose a new condition for topologies named *extensible joint-connectivity* as follows:

(A1) There exist constants $\delta \geq 0$ and $c \geq 1$ and an infinite sequence $1 = t_1 < t_2 < t_3 < \dots$ such that $t_k \leq t_{k-1} + ct_{k-1}^\delta$ and $\cup_{t_{k-1} \leq t < t_k} \mathcal{G}(t)$ is strongly connected for all $k > 1$.

For (A1), we call δ the *extensible exponent* for the joint-connectivity of the topologies. We see for any δ , (A1) is stronger than the infinite joint-connectivity assumption, which can be formulated by $\cup_{t \geq k} \mathcal{G}(t)$ is strongly connected for all $k \geq 1$. On the other hand, for any positive δ (A1) is weaker than the uniform joint-connectivity assumption. In fact, the uniform joint-connectivity is a special case of (A1) with $\delta = 0$.

The extensible joint-connectivity has a possible application to the systems with complex random network topologies, such as the mobile wireless sensor networks or multi robot systems located in the complicated terrain, where the communication probability between two agents depends on their distance and the terrain. In this case the network topologies are a non stationary and strongly correlated random sequence, however it may satisfy (A1) with a probability arbitrarily closing to 1 by choosing suitable c and δ , see the following Corollary 5.1. From this (A1) can increase the robustness of the topology condition greatly compared to the uniform joint-connectivity.

3.2 Sufficient Conditions for Consensus

We firstly give a key lemma deduced from [30]. Before the statement of this key lemma some definitions are needed. For protocol (2.3), set

$$\Phi(t, i) := [I - a(t)L(t)] [I - a(t-1)L(t-1)] \cdots [I - a(i)L(i)].$$

Take $\prod_{l=i}^t (\cdot) := I$ when $t < i$. For any $x \in \mathbb{R}^n$, set $x_{\text{ave}} := \frac{1}{n} \sum_{i=1}^n x_i$ is the average value of x , and define

$$V(x) := \|x - x_{\text{ave}} \mathbf{1}\|^2 = \sum_{i=1}^n (x_i - x_{\text{ave}})^2.$$

For an integer sequence $\{t_k\}_{k \geq 1}$, let

$$k^i := \min\{k : t_k \geq i + 1\} \text{ and } \tilde{k}^t := \max\{k : t_k - 1 \leq t\}.$$

Set

$$d_{\max} := \max_{1 \leq i \leq n} \sum_{j \neq i} a_{ij} \leq (n-1)a_{\max}.$$

Also, following many previous papers[31, 39, 40] we use the balanced assumption for the topology $\mathcal{G}(t)$ of system (2.1)-(2.2):

(A2) The topology $\mathcal{G}(t)$ is balanced for all $t \geq 1$.

Lemma 3.1 Assume (A1) and (A2) are satisfied. Let $z(t) = \Phi(t, i+1)z(i)$ for $t > i$. If $a(t) \in (0, 1/d_{\max})$ then

$$V(z(t)) \leq V(z(i)) \prod_{l=k^i}^{\tilde{k}^t-1} \left(1 - \frac{\delta_l(1-\delta_l)^2\varepsilon_l}{n(n-1)^2}\right),$$

where $\delta_l = \min_{i \neq j} a_{ij} \cdot \min_{t_l \leq t < t_{l+1}} a(t)$ and $\varepsilon_l = \min_{t_l \leq t < t_{l+1}} (1 - a(t)d_{\max})$.

Proof We recall that $L(t)$ is the Laplacian matrix of $\mathcal{G}(t)$. With the assumption of $\mathcal{G}(t)$ is balanced, it is easy to obtain $\mathbf{1}'L(t) = 0$. Take $A(t) = I - a(t)L(t)$. Using the condition of $a(t) \in (0, 1/d_{\max})$, we have $A(t)$ is a double stochastic matrix because $A(t) \geq 0$, $A(t)\mathbf{1} = \mathbf{1}$ and $\mathbf{1}'A(t) = \mathbf{1}'$. Also, we can compute

$$(A'_t A_t)_{ij} \geq (1 - a(t)d_{\max}) [(A_t)_{ij} + (A_t)_{ji}], \quad \forall t \geq 1,$$

and

$$\min_{\emptyset \subset S \subset \{1, \dots, r\}} \sum_{i \in S, j \in S^c} \sum_{t=t_l}^{t_{l+1}-1} [(A_t)_{ij} + (A_t)_{ji}] \geq \min_{i \neq j} a_{ij} \cdot \min_{t_l \leq t < t_{l+1}} a(t) = \delta_l$$

from the condition of $\cup_{t_l \leq t < t_{l+1}} \mathcal{G}(t)$ is strongly connected. With these it is deduced directly from the proof of Theorem 6 in [30] that

$$\begin{aligned} V(z(t_{\tilde{k}^t} - 1)) &= V\left(\Phi(t_{\tilde{k}^t} - 1, t_{\tilde{k}^t-1}) \cdots \Phi(t_{k^i+1} - 1, t_{k^i}) z(t_{k^i} - 1)\right) \\ &\leq V(z(t_{k^i} - 1)) \prod_{l=k^i}^{\tilde{k}^t-1} \left(1 - \frac{\delta_l(1-\delta_l)^2\varepsilon_l}{n(n-1)^2}\right). \end{aligned} \quad (3.1)$$

By Theorem 5 in [30], $V(z(t)) \leq V(z(i))$ for all $t \geq i$, so from (3.1) we can get

$$V(z(t)) \leq V(z(t_{\tilde{k}^t} - 1)) \leq V(z(i)) \prod_{l=k^i}^{\tilde{k}^t-1} \left(1 - \frac{\delta_l(1-\delta_l)^2\varepsilon_l}{n(n-1)^2}\right).$$

□

We also characterize robustness of the protocol (2.1)-(2.2) with respect to the noises. This will be accomplished by accommodating a large class of noises as specified below.

For any random variables X and Y , let $\text{Corr}(X, Y) := \frac{EXY - EXEY}{\sqrt{\text{Var}X \text{Var}Y}}$ denote the linear correlation coefficient between X and Y . Following [49], we employ the notion of $\tilde{\rho}$ -mixing sequence of random variables. Let $\{X_i\}_{i \geq 1}$ be a random variable sequence. For any subset $S, T \subset \mathbb{N}$, write the sub σ -algebras $\mathcal{F}_S := \sigma(X_i, i \in S)$ and set

$$\rho(\mathcal{F}_S, \mathcal{F}_T) := \sup \{ \text{Corr}(X, Y) : X \in L_2(\mathcal{F}_S), Y \in L_2(\mathcal{F}_T) \}.$$

Define the $\tilde{\rho}$ -mixing coefficients by

$$\tilde{\rho}(n) := \sup \left\{ \rho(\mathcal{F}_S, \mathcal{F}_T) : \text{finite sets } S, T \subset \mathbb{N} \text{ such that } \min_{i \in S, j \in T} |i - j| \geq n \right\}$$

for any integer $n \geq 0$. By definition, $0 \leq \tilde{\rho}(n+1) \leq \tilde{\rho}(n) \leq 1$ for all $n \geq 0$, and $\tilde{\rho}(0) = 1$ except for the special case when all X_i are degenerate.

Definition 3.1 A sequence of random variables $\{X_i\}_{i \geq 1}$ is said to be a $\tilde{\rho}$ -mixing sequence if there exists an integer $n > 0$ such that $\tilde{\rho}(n) < 1$.

Under this definition, we give the following assumption for the protocol (2.1)-(2.2):

(A3) The noises $w_{ji}(t), t \geq 1, i = 1, \dots, n, j \in \mathcal{N}_i(t)$ is a $\tilde{\rho}$ -mixing sequence satisfying: (i) $\sup_{i,j,t} \text{Var}(w_{ji}(t)) < \infty$; (ii) there exists a constant ε such that $\max_{1 \leq i \leq n, j \in \mathcal{N}_i(t)} |E[w_{ji}(t)]| = O(\frac{1}{t^\delta \log^\varepsilon t})$, where δ is the same constant appearing in (A1).

Remark 3.1 It is well known that $\tilde{\rho}$ -mixing noises include the ϕ -mixing noises [43, 44], and of course the i.i.d. noise and martingale difference noises; see [50].

A basic property of $\tilde{\rho}$ -mixing sequences is cited here.

Lemma 3.2 (Theorem 2.1 in [51]) Suppose that for an integer $N \geq 1$ and real numbers $q \geq 2$ and $0 \leq r < 1$, there is a positive constant $D = D(q, N, r)$ such that if $\{X_i\}_{i \geq 1}$ is a sequence of random variables with $\tilde{\rho}(N) \leq r$, with $EX_i = 0$ and $E|X_i|^q < \infty$ for every $i \geq 1$. Then for all $n \geq 1$,

$$E \max_{1 \leq i \leq n} \left| \sum_{j=1}^i X_j \right|^q \leq D \left(\sum_{i=1}^n E|X_i|^q + \left(\sum_{i=1}^n E|X_i|^2 \right)^{q/2} \right).$$

Before the statement of our result, we need to give the following three lemmas firstly.

Lemma 3.3 If (A1) is satisfied with $\delta < 1$, then $t_k \leq (ck)^{1/(1-\delta)}$ for all $k \geq 1$.

Lemma 3.4 Let $y := \sum_{t=2}^{\infty} \frac{1}{t \log^\gamma(t)}$ where γ is a constant. Then $y = \infty$ if $\gamma \leq 1$, and $y < \infty$ otherwise.

Lemma 3.5 Suppose (A1) is satisfied with $\delta \leq 1/2$. Let $f(t) = \frac{1}{t^\beta \log^\gamma(t+1)}$ with $\beta \in [\delta, 1-\delta]$ and $\gamma \in (0, 1]$. Then

$$\lim_{t \rightarrow \infty} \sum_{i=1}^t f^2(i) \prod_{j=k^i}^{\tilde{k}^t-1} (1 - c_2 f(t_{j+1})) = 0, \quad (3.2)$$

and

$$\lim_{t \rightarrow \infty} \sum_{i=2}^t \frac{f(i)}{i^\delta \log^\varepsilon i} \prod_{j=k^i}^{\tilde{k}^t-1} (1 - c_2 f(t_{j+1})) = 0, \quad (3.3)$$

where c_2 and ε are two positive constants arbitrarily given.

The proofs of Lemmas 3.3, 3.4, and 3.5 are in the appendix.

The following theorem describes the sufficient conditions for consensus:

Theorem 3.1 Suppose (A1)-(A3) are satisfied with $\delta \leq 1/2$ and $\varepsilon > 0$. Choose $a(t) = \frac{\alpha}{t^{1-\delta} \log^\gamma(t+1)}$, where $\alpha > 0$ and $\gamma \in (\max\{\frac{1}{2}, 1-\varepsilon\}, 1]$ are arbitrary constants. Then for any initial state $x(1)$, system (2.1)-(2.2) reaches consensus in mean square.

Proof By the choice of $a(t)$, there exist an integer $t^* > 1$ such that $a(t)d_{\max} < \frac{1}{2}$ for $t \geq t^*$. We recall

$$\Phi(t, i) = [I - a(t)L(t)] [I - a(t-1)L(t-1)] \cdots [I - a(i)L(i)]$$

and $\prod_{i=j}^t (\cdot) = I$ for $t < j$. For $t > t^*$, using (2.3) repeatedly we get

$$x(t+1) = \Phi(t, t^*)x(t^*) + \sum_{i=t^*}^t a(i)\Phi(t, i+1)\widehat{w}(i). \quad (3.4)$$

Take $\pi = (\frac{1}{n}, \dots, \frac{1}{n}) \in \mathbb{R}^n$. By (A2) we have $\pi L(t) = 0$, so $\pi\Phi(t, i) = \pi$. Set $\widehat{w}^*(i) := \widehat{w}(i) - E[\widehat{w}(i)]$. Take

$$Y(i) = a(i) [\Phi(t, i+1)\widehat{w}^*(i) - (\pi\widehat{w}^*(i)) \mathbf{1}] \in \mathbb{R}^n, \quad (3.5)$$

and

$$Z(i) = a(i) [\Phi(t, i+1)E\widehat{w}(i) - (\pi E\widehat{w}(i)) \mathbf{1}] \in \mathbb{R}^n, \quad (3.6)$$

then by (3.4) we have

$$\begin{aligned} x(t+1) - x_{\text{ave}}(t+1)\mathbf{1} &= x(t+1) - (\pi x(t+1))\mathbf{1} \\ &= \Phi(t, t^*)x(t^*) - (\pi x(t^*))\mathbf{1} + \sum_{i=t^*}^t a(i) [\Phi(t, i+1)\widehat{w}(i) - (\pi\widehat{w}(i))\mathbf{1}] \\ &= \Phi(t, t^*)x(t^*) - (\pi x(t^*))\mathbf{1} + \sum_{i=t^*}^t (Y(i) + Z(i)). \end{aligned} \quad (3.7)$$

Because $EY(i) = \mathbf{0}$, so (3.7) is followed by

$$\begin{aligned} E[V(x(t+1))] &= E\|x(t+1) - x_{\text{ave}}(t+1)\mathbf{1}\|^2 \\ &= \left\| \Phi(t, t^*)x(t^*) - (\pi x(t^*))\mathbf{1} + \sum_{i=t^*}^t Z(i) \right\|^2 + E\left\| \sum_{i=t^*}^t Y(i) \right\|^2 \\ &\leq 2\left\| \Phi(t, t^*)x(t^*) - (\pi x(t^*))\mathbf{1} \right\|^2 + 2\left\| \sum_{i=t^*}^t Z(i) \right\|^2 + E\left\| \sum_{i=t^*}^t Y(i) \right\|^2 \\ &\leq 2V(\Phi(t, t^*)x(t^*)) + 2\left(\sum_{i=t^*}^t \|Z(i)\| \right)^2 + E\left\| \sum_{i=t^*}^t Y(i) \right\|^2. \end{aligned} \quad (3.8)$$

In the following part we will show the right side of (3.8) converges to 0. Firstly by Lemma 3.1 there exists a constant $c_1 > 0$ such that

$$V[\Phi(t, t^*)x(t^*)] \leq V[x(t^*)] \prod_{j=k^{t^*-1}}^{\widetilde{k}^{t-1}} (1 - c_1 a(t_{j+1})) \rightarrow 0 \text{ as } t \rightarrow \infty, \quad (3.9)$$

where the limitation to 0 is deduced from (3.2). Also, similar to (3.9) we can get

$$V(\Psi(t, i+1)E[\widehat{w}(i)]) \leq V(E[\widehat{w}(i)]) \prod_{j=k^i}^{\widetilde{k}^{t-1}} (1 - c_1 a(t_{j+1})).$$

Because $\|Z(i)\|^2 = a^2(i)V(\Psi(t, i+1)E[\widehat{w}(i)])$, using this we have

$$\begin{aligned} \sum_{i=t^*}^t \|Z(i)\| &\leq \sum_{i=t^*}^t a(i) \sqrt{V(E[\widehat{w}(i)])} \prod_{j=k^i}^{\widetilde{k}^t-1} (1 - c_1 a(t_{j+1}))^{1/2} \\ &\leq \sum_{i=t^*}^t a(i) \sqrt{V(E[\widehat{w}(i)])} \prod_{j=k^i}^{\widetilde{k}^t-1} \left(1 - \frac{c_1}{3} a(t_{j+1})\right) \\ &= O\left(\sum_{i=t^*}^t \frac{a(i)}{i^\delta \log^\varepsilon i} \prod_{j=k^i}^{\widetilde{k}^t-1} \left(1 - \frac{c_1}{3} a(t_{j+1})\right)\right), \end{aligned}$$

where the last line uses the condition of $\max_{j,i} |E[w_{ji}(t)]| = O(\frac{1}{t^\delta \log^\varepsilon t})$ in (A3). By (3.3) this yields

$$\lim_{t \rightarrow \infty} \sum_{i=t^*}^{t-1} \|Z(i)\| = 0. \quad (3.10)$$

Furthermore, similar to (3.9) again we have

$$E[V(\Phi(t, i+1)\widehat{w}^*(i))] \leq E[V(\widehat{w}^*(i))] \prod_{j=k^i}^{\widetilde{k}^t-1} (1 - c_1 a(t_{j+1})),$$

so with the fact of $\sup_{i,j,t} \text{Var}(w_{ji}(t)) < \infty$ and (3.2), we obtain

$$\lim_{t \rightarrow \infty} \sum_{i=t^*}^t a^2(i) E[V(\Psi(t, i+1)\widehat{w}^*(i))] = 0.$$

With Lemma 3.2 and this we have

$$\begin{aligned} E\left\|\sum_{i=t^*}^t Y(i)\right\|^2 &= E\sum_{j=1}^n \left[\sum_{i=t^*}^t Y_j(i)\right]^2 \leq 2D \sum_{j=1}^n \sum_{i=t^*}^t EY_j^2(i) \\ &= 2D \sum_{i=t^*}^t E\|Y(i)\|^2 \leq 2D \sum_{i=t^*}^t a^2(i) E[V(\Psi(t, i+1)\widehat{w}^*(i))] \rightarrow 0 \text{ as } t \rightarrow \infty, \end{aligned} \quad (3.11)$$

where D is the same constant appearing in Lemma 3.2. Taking (3.9), (3.10) and (3.11) into (3.8) we get $\lim_{t \rightarrow \infty} E[V(x(t))] = 0$.

It remains to evaluate the limitation of $x_{\text{ave}}(t)$. Let

$$x^* = x_{\text{ave}}(\infty) = \pi x(\infty) = \pi x(1) + \sum_{i=1}^{\infty} a(i) \pi \widehat{w}(i). \quad (3.12)$$

By Lemma 3.4 we obtain

$$\begin{aligned} |Ex^*| &\leq |\pi x(1)| + \sum_{i=1}^{\infty} a(i) |\pi E\widehat{w}(i)| \\ &= |\pi x(1)| + a(1) |\pi E\widehat{w}(1)| + \sum_{i=2}^{\infty} \frac{\alpha}{i^{1-\delta} \log^\gamma(i+1)} \cdot O\left(\frac{1}{i^\delta \log^\varepsilon i}\right) < \infty, \end{aligned} \quad (3.13)$$

and by Lemmas 3.2 and 3.4 we have

$$\begin{aligned}\text{Var}(x^*) &= E \left| \sum_{i=1}^{\infty} a(i) \pi \hat{w}^*(i) \right|^2 = O \left(\sum_{i=1}^{\infty} a^2(i) E |\pi \hat{w}^*(i)|^2 \right) \\ &= O \left(\sum_{i=1}^{\infty} \frac{1}{i^{2(1-\delta)} \log^{2\gamma}(i+1)} \right) < \infty.\end{aligned}\tag{3.14}$$

These together with (3.8) and the fact of $\lim_{t \rightarrow \infty} E[V(x(t))] = 0$ yield that $x(t)$ converges to $x^* \mathbf{1}$ in mean square. \square

Specially, for the zero-mean noises, we can get the following result:

Corollary 3.1 *Suppose (A1)-(A2) are satisfied with $\delta \leq 1/2$. Assume the noises $\{w_{ji}(t)\}$ is a $\tilde{\rho}$ -mixing sequence satisfying $E[w_{ji}(t)] = 0$ for any j, i, t and $\sup_{i,j,t} \text{Var}(w_{ji}(t)) < \infty$. Choose $a(t) = \frac{\alpha}{\sqrt{t \log^\gamma(t+1)}}$, where $\alpha > 0$ and $\gamma \in (\frac{1}{2}, 1]$ are two constants. Then for any initial state $x(1)$, the system (2.1)-(2.2) will reach consensus in mean square.*

Proof With the same process as the proof of Theorem 3.1 we can still get $\lim_{t \rightarrow \infty} E[V(x(t))] = 0$ under the conditions of this corollary. Also, let x^* be the same value defined by (3.12), we have $E[x^*] = \pi x(1)$, and with (3.14) and Lemma 3.4 we have $\text{Var}(x^*) < \infty$. \square

3.3 Necessary Conditions for Consensus

In this subsection we assume $a_{ij} = 1$ for all $i \neq j$. Define \mathcal{A}^* by $\mathcal{A}_{ij}^* = 1$ for $1 \leq i \neq j \leq n$ and $\mathcal{A}_{ii}^* = 0$ for $1 \leq i \leq n$. Let $\mathcal{G}_1 = (\mathcal{V}, \mathcal{E}_1, \mathcal{A}^*)$ be an undirected complete graph, which means each vertex can receive the information of all the others. Let $\mathcal{G}_2 = (\mathcal{V}, \mathcal{E}_2, \mathcal{A}^*)$ be the graph that there exists one undirected edge between vertexes 1 and 2, and on other edge exists. The corresponding Laplacian matrices for \mathcal{G}_1 and \mathcal{G}_2 are

$$L_1 = \begin{pmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & \cdots & -1 \\ \cdots & \cdots & \cdots & \cdots \\ -1 & -1 & \cdots & n-1 \end{pmatrix} \in \mathbb{R}^{n \times n},$$

and

$$L_2 = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

respectively. Define $v_1 := n^{-1/2} \mathbf{1} \in \mathbb{R}^n$, $v_2 := \frac{1}{\sqrt{2}}(1, -1, 0, \dots, 0)' \in \mathbb{R}^n$ and

$$v_i := (i^2 - i)^{-1/2} \underbrace{(1, \dots, 1)}_{i-1}, 1 - i, 0, \dots, 0)' \in \mathbb{R}^n$$

for $i \in [3, n]$. It is easy to compute that: $v_i' v_j = 0$ for $i \neq j$; $L_1 v_1 = \mathbf{0}$; $L_1 v_i = n v_i$ for $i \geq 2$; $L_2 v_2 = 2 v_2$ and $L_2 v_i = \mathbf{0}$ for $i \neq 2$. From this we have the following proposition:

Proposition 3.1 Let $P := (v_1, v_2, \dots, v_n) \in \mathbb{R}^{n \times n}$. Then $P'P = I$, $P \text{diag}(0, n, n, \dots, n)P' = L_1$, and $P \text{diag}(0, 2, 0, \dots, 0)P' = L_2$.

The necessary conditions will contain two parts: one part (Theorem 3.2) discusses the consensus cannot be reached when the extensible exponent δ is larger than a critical value $1/2$, where the simplest i.i.d. noises are used; another part (Theorem 3.3) uses the $\tilde{\rho}$ -mixing noises and shows that if its expectation exceeds a critical value then the consensus cannot be reached.

(A4) The noises $w_{ji}(t), t \geq 1, i = 1, \dots, n, j \in \mathcal{N}_i(t)$ are an i.i.d. random variable sequence whose expectation is 0 and variance is a constant $v > 0$.

Theorem 3.2 Suppose the noises satisfy (A4). Then for any constant $\delta^* > 1/2$, any inconsistent initial state and any gain sequence $\{a(t)\}_{t \geq 1}$, there exist at least one topology sequence $\{\mathcal{G}(t)\}_{t \geq 1} = \{(\mathcal{V}, \mathcal{E}(t), \mathcal{A})\}_{t \geq 1}$ satisfying (A1)-(A2) with $\delta = \delta^*$, such that system (2.1)-(2.2) cannot reach consensus in mean square.

Proof We will show the system (2.1)-(2.2) cannot reach consensus in mean square by contradiction. Because to reach consensus $a(t)$ must converge to 0, there exists an integer k' such that $a(t) < \frac{1}{2n}$ for all $t \geq t_{k'}$. We take $t^* = t_{k'}$. As same as (3.4) we get

$$x(t+1) = \Phi(t, t^*)x(t^*) + \sum_{i=t^*}^t a(i)\Phi(t, i+1)\hat{w}(i). \quad (3.15)$$

We choose $t_k = t_{k-1} + c \lfloor t_{k-1}^\delta \rfloor$ and select

$$\mathcal{G}(t) := \begin{cases} \mathcal{G}_1, & \text{if } t \in \cup_{k=1}^\infty \{t_k^*\}, \\ \mathcal{G}_2, & \text{otherwise,} \end{cases}$$

where

$$t_k^* := \arg \min_{t_k \leq t < t_{k+1}} a(t).$$

Let $\pi = \frac{1}{n}\mathbb{1}'$, then we can compute $\pi L(t) = \mathbf{0}$. So our choice satisfies both (A1) and (A2).

Set $S := \cup_{k=1}^\infty \{t_k^*\}$ and let

$$b_i^t := \prod_{j \in S \cap [i, t]} (1 - na(j)) \quad \text{and} \quad c_i^t := \prod_{j \in S^c \cap [i, t]} (1 - 2a(j)), \quad (3.16)$$

by Proposition 3.1 we can get that

$$\Phi(t, i) = P \text{diag} (1, b_i^t c_i^t, b_i^t, \dots, b_i^t) P'.$$

Set

$$\tilde{\Phi}(t, i) := P \text{diag} (0, b_i^t c_i^t, b_i^t, \dots, b_i^t) P' \quad (3.17)$$

and $x_{\text{ave}}(t) := \frac{1}{n} \sum_{i=1}^n x_i(t)$ be the average value of $x(t)$, by (3.15) we can get

$$\begin{aligned} x(t+1) - x_{\text{ave}}(t+1)\mathbb{1} &= x(t+1) - (\pi x(t+1))\mathbb{1} \\ &= \Phi(t, t^*)x(t^*) - (\pi x(t^*))\mathbb{1} + \sum_{i=t^*}^{t-1} a(i) [\Phi(t, i+1)\hat{w}(i) - (\pi \hat{w}(i))\mathbb{1}] \\ &\quad + a(t) [\hat{w}(t) - (\pi \hat{w}(t))\mathbb{1}] \\ &= \tilde{\Phi}(t, t^*)x(t^*) + \sum_{i=t^*}^{t-1} a(i)\tilde{\Phi}(t, i+1)\hat{w}(i) + a(t) [\hat{w}(t) - (\pi \hat{w}(t))\mathbb{1}]. \end{aligned} \quad (3.18)$$

Because $\{w_{ji}(t)\}$ is a sequence of i.i.d. random variables with $Ew_{ji}(t) = 0$,

$$\begin{aligned} E[V(x(t+1))] &= E\|\tilde{\Phi}(t, t^*)x(t^*)\|^2 + \sum_{i=t^*}^{t-1} a^2(i)E\|\tilde{\Phi}(t, i+1)\hat{w}(i)\|^2 \\ &\quad + a^2(t)E\|\hat{w}(t) - (\pi\hat{w}(t))\mathbf{1}\|^2 \\ &\geq \sum_{i=t^*}^{t-1} a^2(i)E\|\tilde{\Phi}(t, i+1)\hat{w}(i)\|^2. \end{aligned} \quad (3.19)$$

Also, for any $y = \tilde{\Phi}(t, i)x$ we can compute

$$y_1 = \frac{x_1 - x_2}{2} b_i^t c_i^t + \left(\frac{x_1 + x_2}{2} - \pi x \right) b_i^t. \quad (3.20)$$

Substituting this into (3.19), we get

$$\begin{aligned} E[V(x(t+1))] &\geq \sum_{i=t^*}^{t-1} a^2(i)(b_{i+1}^t)^2 E \left[\frac{\hat{w}_1(i) - \hat{w}_2(i)}{2} c_{i+1}^t + \frac{\hat{w}_1(i) + \hat{w}_2(i)}{2} - \pi \hat{w}(i) \right]^2 \\ &\geq \sum_{i=t^*}^{t-1} a^2(i)(b_{i+1}^t)^2 \left(\frac{1}{2} - \frac{1}{n} \right)^2 E \hat{w}_1^2(i) \geq \left(\frac{1}{2} - \frac{1}{n} \right)^2 v \sum_{i=t^*}^{t-1} a^2(i)(b_{i+1}^t)^2. \end{aligned} \quad (3.21)$$

To reach consensus the last line of (3.21) should converge to 0. We show that this is impossible.

First, because the initial state $x(1)$ is not consistent, to reach consensus we must choose some positive $a(t)$. Considering the affection of noises, we must select positive $a(t)$ infinite times to guarantee consensus in mean square. Thus, we can pick $t' \geq t^*$ such that $a(t') > 0$. If the last line of (3.21) converge to 0, we have

$$\prod_{k=k'}^{\infty} (1 - na(t_k^*)) = \lim_{t \rightarrow \infty} b_{t^*}^t \leq \lim_{t \rightarrow \infty} b_{t'+1}^t = 0. \quad (3.22)$$

Also, we recall that $t_{k^t} \leq t + 1$, so

$$\begin{aligned} \sum_{i=t^*}^t a^2(i)(b_{i+1}^{t+1})^2 &\geq \sum_{k=k'}^{\tilde{k}^t-1} \sum_{i=t_k}^{t_{k+1}-1} a^2(i)(b_{i+1}^{t+1})^2 \\ &\geq \sum_{k=k'}^{\tilde{k}^t-1} \sum_{i=t_k}^{t_{k+1}-1} a^2(i) \prod_{j=k}^{\tilde{k}^t} [1 - na(t_j^*)]^2 \\ &\geq \sum_{k=k'}^{\tilde{k}^t-1} (t_{k+1} - t_k) a^2(t_k^*) \prod_{j=k}^{\tilde{k}^t} [1 - na(t_j^*)]^2 \\ &> \frac{1}{16} \sum_{k=k'}^{\tilde{k}^t-1} \lfloor ct_k^\delta \rfloor a^2(t_k^*) \prod_{j=k+1}^{\tilde{k}^t-1} [1 - na(t_j^*)]^2 \\ &\geq \frac{1}{16} \sum_{k=k'}^{\tilde{k}^t-1} \lfloor ct_k^\delta \rfloor a^2(t_k^*) \prod_{j=k+1}^{\tilde{k}^t-1} [1 - 2na(t_j^*)]. \end{aligned} \quad (3.23)$$

Let $I_{\{\cdot\}}$ be the indicator function, then

$$1 - 2na(t_j^*) = \left[1 - 2na(t_j^*)I_{\{a(t_j^*) > t_j^{-\delta}\}}\right] \left[1 - 2na(t_j^*)I_{\{a(t_j^*) \leq t_j^{-\delta}\}}\right]. \quad (3.24)$$

Because $c \geq 1$ and $\delta > 1/2$, by the choice of t_k , we have

$$t_k = t_{k-1} + c \lfloor t_{k-1}^\delta \rfloor \geq t_{k-1} + \frac{c}{2} t_{k-1}^\delta > t_{k-1} + \frac{\sqrt{t_{k-1}}}{2},$$

then by induction we can get $t_k > \frac{1}{20}k^2$. So,

$$\sum_{j=k'}^{\infty} a(t_j^*)I_{\{a(t_j^*) \leq t_j^{-\delta}\}} \leq \sum_{j=k'}^{\infty} t_j^{-\delta} < \sum_{j=k'}^{\infty} 20^\delta j^{-2\delta} < \infty, \quad (3.25)$$

which indicates

$$c_1 := \prod_{j=k'}^{\infty} \left[1 - 2na(t_j^*)I_{\{a(t_j^*) \leq t_j^{-\delta}\}}\right] > 0.$$

Substituting this and (3.24) into (3.23), we have

$$\begin{aligned} \sum_{i=t^*}^t a^2(i)(b_{i+1}^{t+1})^2 &\geq \frac{c_1}{16} \sum_{k=k'}^{\tilde{k}^t-1} \lfloor ct_k^\delta \rfloor a^2(t_k^*) \prod_{j=k+1}^{\tilde{k}^t-1} \left[1 - 2na(t_j^*)I_{\{a(t_j^*) > t_j^{-\delta}\}}\right] \\ &\geq \frac{c_1}{16} \sum_{k=k'}^{\tilde{k}^t-1} \lfloor ct_k^\delta \rfloor a^2(t_k^*) I_{\{a(t_k^*) > t_k^{-\delta}\}} \prod_{j=k+1}^{\tilde{k}^t-1} \left[1 - 2na(t_j^*)I_{\{a(t_j^*) > t_j^{-\delta}\}}\right] \\ &> \frac{c_1}{16} \sum_{k=k'}^{\tilde{k}^t-1} \lfloor ct_k^\delta \rfloor t_k^{-\delta} a(t_k^*) I_{\{a(t_k^*) > t_k^{-\delta}\}} \prod_{j=k+1}^{\tilde{k}^t-1} \left[1 - 2na(t_j^*)I_{\{a(t_j^*) > t_j^{-\delta}\}}\right] \\ &\geq \frac{c_1 c}{32} \sum_{k=k'}^{\tilde{k}^t-1} a(t_k^*) I_{\{a(t_k^*) > t_k^{-\delta}\}} \prod_{j=k+1}^{\tilde{k}^t-1} \left[1 - 2na(t_j^*)I_{\{a(t_j^*) > t_j^{-\delta}\}}\right] \\ &= \frac{c_1 c}{64n} \left(1 - \prod_{j=k'}^{\tilde{k}^t-1} \left[1 - 2na(t_j^*)I_{\{a(t_j^*) > t_j^{-\delta}\}}\right]\right). \end{aligned} \quad (3.26)$$

Also, by (3.22) and (3.25), we have

$$\sum_{j=k'}^{\infty} a(t_j^*)I_{\{a(t_j^*) > t_j^{-\delta}\}} = \infty,$$

so by (3.26) we can get

$$\liminf_{t \rightarrow \infty} \sum_{i=t^*}^t a^2(i)(b_{i+1}^{t+1})^2 \geq \frac{c_1 c}{64n}.$$

Combining this with (3.21) we see the system cannot reach consensus in mean square. \square

Theorem 3.3 Suppose (A1)-(A3) are satisfied with $\varepsilon \leq 0$. Then for any inconsistent initial state and any gain sequence $\{a(t)\}_{t \geq 1}$, there exist at least on topology sequence $\{\mathcal{G}(t)\}_{t \geq 1}$ and one noise sequence $\{w_{ji}(t)\}_{t \geq 1, j \in \mathcal{N}_i(t)}$ such that system (2.1)-(2.2) cannot reach consensus in mean square.

The proof of this theorem mainly uses the idea of the proof of Theorem 3.2 and is put in Appendix.

4 Fastest Convergence Rates of Consensus

This section ascertains the fastest rate of our system converging to consensus among all the gain functions under the unknown switching topologies. Some lower bounds of the fastest convergence rate is given in Subsection 4.1, and an upper bound is given in Subsection 4.2.

4.1 lower bounds

Differing from the noise-free system[23, 24, 47], the limitation value x^* in Definition 2.1 is unknown. Also, if the system (2.1)-(2.2) reaches consensus in mean square, it must be true that $\lim_{t \rightarrow \infty} E[V(x(t))] = 0$, so we use $E[V(x(t))]$ to measure the convergence rate of consensus instead of $E\|x(t) - x^*\|^2$. In this paper the *fastest convergence rate* of consensus at time t is the minimal value of $E[V(x(t))]$ among all the controls $a(1) \geq 0, a(2) \geq 0, \dots, a(t-1) \geq 0$. Because this rate is very complex which depends on the unknown time-varying topologies, the noises and the initial state, we need introduce some other definitions to simplify its expression. First we define

$$\rho_1(t) := \inf_{a(1) \geq 0, a(2) \geq 0, \dots, a(t-1) \geq 0} \inf_{\mathcal{G}(1), \mathcal{G}(2), \dots, \mathcal{G}(t-1)} E[V(x(t))] \quad (4.1)$$

to be the fastest convergence rate concerning with the best topologies. Here we recall that $\{\mathcal{G}(t)\}_{t \geq 1} = \{(\mathcal{V}, \mathcal{E}(t), \mathcal{A})\}_{t \geq 1}$ is the topology sequence and note that $\rho_1(t)$ depends on the noises and the initial state $x(1)$.

We also consider the fastest convergence rate concerning with the worst topologies. Define $\mathcal{G}_{\delta, c}$ to be the set of the topology sequence satisfying (A1)-(A2), where δ, c , are the constants appearing in (A1). Let

$$\rho_2(t) := \inf_{a(1) \geq 0, a(2) \geq 0, \dots, a(t-1) \geq 0} \sup_{\{\mathcal{E}(k)\} \in \mathcal{G}_{\delta, c}} E[V(x(t))] \quad (4.2)$$

denote the fastest convergence rate with respect to the worst topologies satisfying (A1)-(A2). We note that $\rho_2(t)$ depends on δ, c , the noises and the initial state. By the definitions of $\rho_1(t)$ and $\rho_2(t)$ we have for any topology sequence satisfying (A1)-(A2), its fastest convergence rate will be not faster than $\rho_1(t)$ but not slower than $\rho_2(t)$ providing the same initial state and noises.

In the following part of this subsection we will give a lower bound for $\rho_2(t)$ and $\rho_1(t)$ respectively under the i.i.d. noises. The lower bound of the fastest convergence rate indicates whatever the controls are the convergence rate will be not faster than it. Intuitively, a lower bound under the simplest i.i.d. noises is still a lower bound under some other noises.

Theorem 4.1 Suppose (A1)-(A2) and (A4) are satisfied with $\delta < 1/2$. Then for any inconsistent initial state, under protocol (2.1)-(2.2) there exists a constant $C = C(c, \delta) \in (0, 1)$ such that

$$\rho_2(t) \geq \min \left\{ \frac{Cv}{nt^{1-2\delta}}, CV(x(1)), \frac{Cv}{n^2}, \frac{2(n-1)v}{9n^3} \right\}, \quad \forall t \geq 1, \quad (4.3)$$

where δ, c and v are the same constants appearing in (A1) and (A4).

Proof Without loss of generality, we assume

$$\sum_{i=3}^n [x_i(1) - x_{\text{ave}}(1)]^2 \geq \frac{n-2}{n} V(x(1)) > 0. \quad (4.4)$$

Choose a_{ij}, t_k and $\mathcal{G}(t)$ as same as the proof of Theorem 3.2, and take $\pi = \frac{1}{n}\mathbf{1}'$. Also, Using induction and Lemma 3.3 we can get there exists a constant $c_1 := c_1(c, \delta) > 0$ such that

$$c_1 k^{1/(1-\delta)} \leq t_k \leq (ck)^{1/(1-\delta)}, \quad \forall k \geq 1. \quad (4.5)$$

For the time $t+2$ with $t \geq 0$, we consider all the choices of $\{a(i)\}_{i=1}^{t+1}$ to get a lower bound of $E\|x(t+2) - x_{\text{ave}}(t+2)\mathbf{1}\|^2$. Let $t^* \in [1, t+2]$ be the minimum time such that if $t \geq t^*$ then $a(t) < \frac{1}{3n}$. We see if $t^* \geq 2$ then $a(t^*-1) \geq \frac{1}{3n}$. By (3.19) we have

$$\begin{aligned} E[V(x(t+2))] &= E\|\tilde{\Phi}(t+1, t^*)x(t^*)\|^2 + \sum_{i=t^*}^t a^2(i) E\|\tilde{\Phi}(t+1, i+1)\hat{w}(i)\|^2 \\ &\quad + a^2(t+1) E\|\hat{w}(t+1) - (\pi\hat{w}(t+1))\mathbf{1}\|^2, \end{aligned} \quad (4.6)$$

here we assume $t^* < t+2$ because otherwise, by (3.19),

$$\begin{aligned} E[V(x(t+2))] &\geq a^2(t+1) E\|\hat{w}(t+1) - (\pi\hat{w}(t+1))\mathbf{1}\|^2 \\ &\geq a^2(t^*-1) \left(2 - \frac{2}{n}\right) v \geq \frac{2(n-1)v}{9n^3}, \end{aligned} \quad (4.7)$$

which is followed by our result directly.

We recall that $b_i^t := \prod_{j \in S \cap [i, t]} (1 - na(j))$ defined in (3.16). For any $y = \tilde{\Phi}(t, i)x$ we can compute

$$y_j = (x_j - \pi x) b_i^t, \quad \forall 3 \leq j \leq n,$$

so if $t^* = 1$ then

$$\begin{aligned} E\|\tilde{\Phi}(t+1, t^*)x(t^*)\|^2 &\geq (b_{t^*}^{t+1})^2 \sum_{j=3}^n [x_j(1) - x_{\text{ave}}(1)]^2 \\ &\geq \frac{(n-2)(b_{t^*}^{t+1})^2}{n} V(x(1)), \end{aligned} \quad (4.8)$$

where the last inequality uses (4.4). Otherwise, by the choice of t^* we have $a(t^*-1) \geq \frac{1}{3n}$, so

$$\begin{aligned} &E\|\tilde{\Phi}(t+1, t^*)x(t^*)\|^2 \\ &= E\|\tilde{\Phi}(t+1, t^*) \{[I - a(t^*-1)L(t^*-1)]x(t^*-1) + a(t^*-1)\hat{w}(t^*-1)\}\|^2 \\ &\geq a^2(t^*-1) E\|\tilde{\Phi}(t+1, t^*)\hat{w}(t^*-1)\|^2 \geq \frac{v}{9n^2} \left(\frac{1}{2} - \frac{1}{n}\right)^2 (b_{t^*}^{t+1})^2, \end{aligned} \quad (4.9)$$

where the last inequality uses (3.21). Set $k' := k^{t^*-1}$, then we have $t_{k'} \geq t^*$ but $t_{k'-1} < t^*$. Similar to (3.21) and (3.23), we have

$$\begin{aligned}
& \sum_{i=t^*}^t a^2(i) E \|\tilde{\Phi}(t+1, i+1) \widehat{w}(i)\|^2 \\
& \geq \left(\frac{1}{2} - \frac{1}{n}\right)^2 v \sum_{i=t^*}^t a^2(i) (b_{i+1}^{t+1})^2 \\
& \geq \frac{4(n-2)^2 v}{81n^2} \sum_{k=k'}^{\tilde{k}^t-1} \lfloor ct_k^\delta \rfloor a^2(t_k^*) \prod_{j=k+1}^{\tilde{k}^t-1} [1 - 2na(t_j^*)].
\end{aligned} \tag{4.10}$$

Similar to (3.24) we have

$$1 - 2na(t_j^*) = \left[1 - 2na(t_j^*) I_{\{a(t_j^*) > t^{2\delta-1}/\lfloor ct_j^\delta \rfloor\}}\right] \left[1 - 2na(t_j^*) I_{\{a(t_j^*) \leq t^{2\delta-1}/\lfloor ct_j^\delta \rfloor\}}\right].$$

According to (4.5) and the fact $t_{\tilde{k}^t} \leq t+1$, we get

$$\begin{aligned}
\sum_{j=1}^{\tilde{k}^t-1} a(t_j^*) I_{\{a(t_j^*) \leq t^{2\delta-1}/\lfloor ct_j^\delta \rfloor\}} & \leq \sum_{j=1}^{\tilde{k}^t-1} \frac{t^{2\delta-1}}{\lfloor ct_j^\delta \rfloor} = t^{2\delta-1} \sum_{j=1}^{\tilde{k}^t-1} O\left(j^{\frac{-\delta}{1-\delta}}\right) \\
& = t^{2\delta-1} O\left((\tilde{k}^t)^{\frac{1-2\delta}{1-\delta}}\right) \\
& = O\left(t_{\tilde{k}^t}^{2\delta-1} (\tilde{k}^t)^{\frac{1-2\delta}{1-\delta}}\right) = O\left((\tilde{k}^t)^{\frac{2\delta-1}{1-\delta}} (\tilde{k}^t)^{\frac{1-2\delta}{1-\delta}}\right) = O(1),
\end{aligned}$$

so with the fact of $a(j) < \frac{1}{3n}$ for $j \geq t^*$,

$$\begin{aligned}
& \prod_{j=k'}^{\tilde{k}^t-1} \left[1 - 2na(t_j^*) I_{\{a(t_j^*) \leq t^{2\delta-1}/\lfloor ct_j^\delta \rfloor\}}\right] \\
& \geq \inf_{t \geq 0} \prod_{j=1}^{\tilde{k}^t-1} \left[1 - \min\left\{\frac{2n}{\lfloor ct_j^\delta \rfloor t^{1-2\delta}}, \frac{2}{3}\right\}\right] := c_2 > 0.
\end{aligned}$$

Here we use the setting of $\prod_{j=i}^k (\cdot) = 1$ for $k < i$. Similar to (3.26) we have

$$\begin{aligned}
& \sum_{k=k'}^{\tilde{k}^t-1} \lfloor ct_k^\delta \rfloor a^2(t_k^*) \prod_{j=k+1}^{\tilde{k}^t-1} [1 - 2na(t_j^*)] \\
& \geq \frac{c_2}{t^{1-2\delta}} \sum_{k=k'}^{\tilde{k}^t-1} a(t_k^*) I_{\{a(t_j^*) > t^{2\delta-1}/\lfloor ct_j^\delta \rfloor\}} \prod_{j=k+1}^{\tilde{k}^t-1} \left[1 - 2na(t_j^*) I_{\{a(t_j^*) > t^{2\delta-1}/\lfloor ct_j^\delta \rfloor\}}\right] \\
& = \frac{c_2}{2nt^{1-2\delta}} \left(1 - \prod_{j=k'}^{\tilde{k}^t-1} \left[1 - 2na(t_j^*) I_{\{a(t_j^*) > t^{2\delta-1}/\lfloor ct_j^\delta \rfloor\}}\right]\right).
\end{aligned} \tag{4.11}$$

If

$$\prod_{j=k'}^{\tilde{k}^t-1} \left[1 - 2na(t_j^*) I_{\{a(t_j^*) > t^{2\delta-1}/\lfloor ct_j^\delta \rfloor\}}\right] \leq \frac{1}{2},$$

then together (4.11), (4.10) and (4.6) our result is obtained. Otherwise,

$$\begin{aligned} b_{t^*}^{t+1} &\geq c_2 \prod_{j=k'-1}^{\tilde{k}^t+1} \left[1 - na(t_j^*) I_{\{a(t_j^*) > t^{2\delta-1}/\lfloor ct_j^\delta \rfloor\}} \right] \\ &\geq c_2 \left(\frac{2}{3} \right)^3 \prod_{j=k'}^{\tilde{k}^t-1} \left[1 - na(t_j^*) I_{\{a(t_j^*) > t^{2\delta-1}/\lfloor ct_j^\delta \rfloor\}} \right] > \left(\frac{2}{3} \right)^3 \frac{c_2}{2}. \end{aligned}$$

Substituting this into (4.8) and (4.9), the desired result follows. \square

Before the estimation of $\rho_1(t)$ we need introduce the following lemma:

Lemma 4.1 *Let $L \in \mathbb{R}^{n \times n}$ be the Laplacian matrix of a weighted directed graph arbitrarily given. For any $x \in \mathbb{R}^n$, let $y := (I - aL)x$, where $a > 0$ is a constant. Then*

$$V(y) \geq (1 - a\lambda_{\max}(L + L')) V(x),$$

where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue.

The proof of this lemma is in the appendix.

The following theorem gives a lower bound of $\rho_1(t)$ under the i.i.d. noises.

Theorem 4.2 *Suppose the noises assume (A4). Then for any inconsistent initial state, under the protocol (2.1)-(2.2) we have*

$$\rho_1(t) \geq \min \left\{ \frac{1}{e} V(x(1)), \frac{v}{16en(n-1)a_{\max}^2 t} \right\}, \quad \forall t \geq 1.$$

Proof For any $t > 1$, any control gains $\{a(k)\}_{k=1}^t$ and any topologies $\{\mathcal{G}(k)\}_{k=1}^t$, we consider the following two cases:

Case I: The edge set of $\mathcal{G}(k)$ is not empty for all $1 \leq k \leq t$. This indicates

$$E[V(\hat{w}(k))] \geq \frac{(n-1)v}{n} \min_{i \neq j} a_{ij}^2 \geq \frac{(n-1)v}{n} := c_1. \quad (4.12)$$

Here we use the fact of $1 \leq a_{ij} \leq a_{\max}$ for $i \neq j$. To simplify the exposition we take $c_2 = 4(n-1)a_{\max}$. Let $t^* \in [1, t+1]$ be the minimum time such that if $t \geq t^*$ then $a(t) < \frac{1}{c_2}$. If $t^* = t+1$, we have $a(t) \geq \frac{1}{c_2}$, then similar to (4.7),

$$E[V(x(t+1))] \geq a^2(t) E[V(\hat{w}(t))] \geq \frac{c_1}{c_2^2} = \frac{v}{16n(n-1)a_{\max}^2},$$

which is followed by our result directly.

In the rest part of this case we assume $t^* \leq t$. Similar to (4.6) we can get

$$E[V(x(t+1))] = E[V(\Phi(t, t^*)x(t^*))] + \sum_{i=t^*}^t a^2(i) E[V(\Phi(t, i+1)\hat{w}(i))]. \quad (4.13)$$

Also, by Gershgorin circle theorem,

$$\begin{aligned} \lambda_{\max}(L(t) + L'(t)) &\leq \max_{1 \leq i \leq n} \left(2L_{ii}(t) + \sum_{j \neq i} |L_{ji}(t) + L_{ij}(t)| \right) \\ &\leq 4(n-1)a_{\max} = c_2, \quad \forall t \geq 1. \end{aligned}$$

Using Lemma 4.1 and this, (4.13) yields

$$\begin{aligned} E[V(x(t+1))] &\geq E[V(x(t^*))] \prod_{j=t^*}^t (1 - c_2 a(j)) \\ &\quad + \sum_{i=t^*}^t a^2(i) E[V(\widehat{w}(i))] \prod_{j=i+1}^t (1 - c_2 a(j)). \end{aligned} \quad (4.14)$$

Let $I_{\{\cdot\}}$ be the indicator function. Because

$$1 - c_2 a(j) = \left[1 - c_2 a(j) I_{\{a(j) > \frac{1}{c_2 t}\}} \right] \left[1 - c_2 a(j) I_{\{a(j) \leq \frac{1}{c_2 t}\}} \right],$$

and

$$\prod_{j=1}^t \left(1 - c_2 a(j) I_{\{a(j) \leq \frac{1}{c_2 t}\}} \right) \geq \left(1 - \frac{1}{t} \right)^t \geq \frac{1}{e},$$

from (4.14) and (4.12), we obtain

$$\begin{aligned} E[V(x(t+1))] &\geq \frac{1}{e} E[V(x(t^*))] \prod_{j=t^*}^t \left(1 - c_2 a(j) I_{\{a(j) > \frac{1}{c_2 t}\}} \right) \\ &\quad + \frac{c_1}{e} \sum_{i=t^*}^t a^2(i) \prod_{j=i+1}^t \left(1 - c_2 a(j) I_{\{a(j) > \frac{1}{c_2 t}\}} \right). \end{aligned} \quad (4.15)$$

It remains to discuss the value of the right side of (4.15). We firstly consider $E[V(x(t^*))]$. If $t^* = 1$ then $E[V(x(t^*))] = V(x(1))$. Otherwise, by the definition of t^* we have $a(t^* - 1) \geq \frac{1}{c_2}$, so similar to (4.13) and by (4.12) we have

$$E[V(x(t^*))] \geq a^2(t^* - 1) E[V(\widehat{w}(t^* - 1))] \geq \frac{c_1}{c_2^2}.$$

Thus, together these discussion we have

$$E[V(x(t^*))] \geq \min \left\{ V(x(1)), \frac{c_1}{c_2^2} \right\}. \quad (4.16)$$

Also, because $y_1 \geq y_2 I_{\{y_1 > y_2\}}$ for any $y_1, y_2 \geq 0$, we can get

$$\begin{aligned} &\sum_{i=t^*}^t a^2(i) \prod_{j=i+1}^t \left(1 - c_2 a(j) I_{\{a(j) > \frac{1}{c_2 t}\}} \right) \\ &\geq \sum_{i=t^*}^t \frac{a(i)}{c_2 t} I_{\{a(i) > \frac{1}{c_2 t}\}} \prod_{j=i+1}^t \left(1 - c_2 a(j) I_{\{a(j) > \frac{1}{c_2 t}\}} \right) \\ &= \frac{1}{c_2^2 t} \left(1 - \prod_{j=t^*}^t \left(1 - c_2 a(j) I_{\{a(j) > \frac{1}{c_2 t}\}} \right) \right). \end{aligned} \quad (4.17)$$

Because $by_1 + (1 - b)y_2 \geq \min\{y_1, y_2\}$ for any $b \in [0, 1]$, substituting (4.17) into (4.15) we have

$$\begin{aligned} E[V(x(t+1))] &\geq \min \left\{ \frac{1}{e} E[V(x(t^*))], \frac{c_1}{ec_2^2 t} \right\} \\ &\geq \min \left\{ \frac{1}{e} V(x(1)), \frac{c_1}{ec_2^2 t} \right\}, \end{aligned}$$

where the last inequality uses (4.16).

Case II: $\{\mathcal{G}(k)\}_{k=1}^t$ contains m graphs whose edge set is empty. By Case I it is easy to get

$$E[V(x(t+1))] \geq \min \left\{ \frac{1}{e} V(x(1)), \frac{c_1}{ec_2^2(t-m)} \right\} \geq \min \left\{ \frac{1}{e} V(x(1)), \frac{c_1}{ec_2^2 t} \right\}.$$

□

4.2 An upper bound

Before providing the statement of our result, we need a lemma. Recall that $k^i = \min\{k : t_k \geq i + 1\}$ and $\tilde{k}^t = \max\{k : t_k - 1 \leq t\}$, we firstly give:

Lemma 4.2 *Suppose (A1) is satisfied with $\delta < 1/2$, then for any constant $c_1 > 0$ and integer $t^* \geq 0$,*

$$\prod_{j=k^i}^{\tilde{k}^t-1} \left(1 - \frac{c_1}{t_{j+1}^{1-\delta} + t^*} \right) < \left(\frac{i^{1-\delta} + 2c + t^*}{(t+1)^{1-\delta} + t^*} \right)^{\frac{c_1}{2c}},$$

where c is the same constant appearing in (A1).

The proof of Lemma 4.2 is in the appendix.

In this subsection we will estimate the upper bound of the fastest convergence rate, which indicates whatever the controls are the convergence rate will be not slower than this bound. Differing from the lower bound, the estimation of upper bound need consider a more complex class of noises than the i.i.d. noises. Similar to (A3) this subsection uses the $\tilde{\rho}$ -mixing noises with the following modifications.

(A5) The noise $w_{ji}(t), t \geq 1, i = 1, \dots, n, j \in \mathcal{N}_i(t)$ is a $\tilde{\rho}$ -mixing sequence satisfying: (i) $v := \sup_{i,j,t} \text{Var}(w_{ji}(t)) < \infty$; (ii) there exists a constant $\overline{C} > 0$ such that $\max_{1 \leq i \leq n, j \in \mathcal{N}_i(t)} |E[w_{ji}(t)]| \leq \overline{C}t^{-1/2}$ for any $t \geq 1$.

Recall that $\mathcal{G}_{\delta,c}$ is the set of the topology sequence $\{\mathcal{G}(t)\}_{t \geq 1}$ satisfying (A1)-(A2), where δ, c , are the constants appearing in (A1). Let \mathcal{W}^* denote the set of the noise sequences satisfying (A5), and let

$$\rho_3(t) := \inf_{a(1) \geq 0, a(2) \geq 0, \dots, a(t-1) \geq 0} \sup_{\{\mathcal{G}(k)\} \in \mathcal{G}_{\delta,c}, \{w_{ji}(k)\} \in \mathcal{W}^*} E[V(x(t))] \quad (4.18)$$

denote the fastest convergence speed with respect to the worst cases of the topologies and noises. If the noises satisfy (A4) then the definition of $\rho_3(t)$ is equivalent to $\rho_2(t)$.

Theorem 4.3 Suppose (A1)-(A2) and (A5) are satisfied with $\delta < 1/2$. Then for any initial state, system (2.1)-(2.2) can reach consensus with a rate of

$$\begin{aligned} \rho_3(t) &< 2V(x(1)) \left(\frac{2c + t^*}{t^{1-\delta} + t^*} \right)^2 \\ &+ \left(Dv + 4\bar{C}^2 \right) \left(1 + \frac{c}{\alpha(n-1)a_{\max}} \right)^2 \frac{2(n-1)a_{\max}^2 \alpha^2 t}{(t^{1-\delta} + t^*)^2}, \quad \forall t \geq 1, \end{aligned} \quad (4.19)$$

where c, v, \bar{C}, D are the same constants appearing in (A1), (A5), and Lemma 3.2, and α and t^* are two constants satisfying

$$\alpha \geq \frac{32n(n-1)^4 c}{(2n-3)^2} \quad \text{and} \quad t^* \geq \lfloor 2\alpha(n-1)a_{\max} \rfloor. \quad (4.20)$$

Proof For any topology sequence satisfies (A1)-(A2) with $\delta < 1/2$ and any noise sequence satisfies (A5), we choose $a(t) = \frac{\alpha}{t^{1-\delta} + t^*}$. By (3.8) we have

$$E[V(x(t+1))] \leq 2V(\Phi(t, 1)x(1)) + 2 \left(\sum_{i=1}^t \|Z(i)\| \right)^2 + E \left\| \sum_{i=1}^t Y(i) \right\|^2. \quad (4.21)$$

where $Y(i)$ and $Z(i)$ are defined by (3.5) and (3.6) respectively. In the following part we estimate the values of $E \left\| \sum_{i=1}^t Y(i) \right\|^2$, $\sum_{i=1}^t \|Z(i)\|$ and $V(\Phi(t, 1)x(1))$ respectively.

Firstly by (3.11),

$$E \left\| \sum_{i=1}^t Y(i) \right\|^2 \leq 2D \sum_{i=1}^t a^2(i) E[V(\Phi(t, i+1)\hat{w}^*(i))]. \quad (4.22)$$

Here we recall $\Phi(t, i+1) = I$ when $t \leq i$. Because $1 \leq a_{ij} \leq a_{\max}$ for all $i \neq j$, and $a(t) < \frac{1}{2(n-1)a_{\max}}$ for all $t \geq 1$, by Lemma 3.1, (4.20) and Lemma 4.2 we have for any $x \in \mathbb{R}^n$,

$$\begin{aligned} E[V(\Phi(t, i+1)x)] &\leq E[V(x)] \prod_{j=k^i}^{\tilde{k}^{t-1}} \left(1 - \left(1 - \frac{1}{2(n-1)} \right)^2 \frac{a(t_{j+1})}{2n(n-1)^2} \right) \\ &\leq E[V(x)] \prod_{j=k^i}^{\tilde{k}^{t-1}} \left(1 - \frac{4c}{t_{j+1}^{1-\delta} + t^*} \right) < E[V(x)] \left(\frac{2c + i^{1-\delta} + t^*}{(t+1)^{1-\delta} + t^*} \right)^2. \end{aligned} \quad (4.23)$$

Using this and (4.22) we can get

$$E \left\| \sum_{i=1}^t Y(i) \right\|^2 \leq 2D \sum_{i=1}^t \frac{\alpha^2}{(i^{1-\delta} + t^*)^2} E[V(\hat{w}^*(i))] \left(\frac{2c + i^{1-\delta} + t^*}{(t+1)^{1-\delta} + t^*} \right)^2. \quad (4.24)$$

By Jensen's inequality we obtain

$$\begin{aligned} E[V(\hat{w}^*(i))] &= \sum_{j=1}^n \frac{1}{n} E[\hat{w}_j^*(i)]^2 - E \left[\sum_{j=1}^n \frac{\hat{w}_j^*(i)}{n} \right]^2 \leq \sum_{j=1}^n \frac{1}{n} E[\hat{w}_j^*(i)]^2 \\ &= \sum_{j=1}^n \frac{1}{n} E \left[\sum_{k \in \mathcal{N}_j(t)} a_{jk} w_{kj}^*(i) \right]^2 \leq \sum_{j=1}^n \frac{1}{n} E \left[\sum_{k \in \mathcal{N}_j(t)} a_{jk} \sum_{k \in \mathcal{N}_j(t)} a_{jk} (w_{kj}^*(i))^2 \right] \\ &\leq \sum_{j=1}^n \frac{1}{n} v(n-1)^2 a_{\max}^2 = v(n-1)^2 a_{\max}^2, \end{aligned} \quad (4.25)$$

where $w_{kj}^*(i) := w_{kj}(i) - Ew_{kj}(i)$. Substituting this into (4.24), we have

$$\begin{aligned} E \left\| \sum_{i=1}^t Y(i) \right\|^2 &\leq \frac{2Dv(n-1)^2 a_{\max}^2 \alpha^2}{[(t+1)^{1-\delta} + t^*]^2} \sum_{i=1}^t \frac{(2c + i^{1-\delta} + t^*)^2}{(i^{1-\delta} + t^*)^2} \\ &\leq \frac{2Dv(n-1)^2 a_{\max}^2 \alpha^2}{[(t+1)^{1-\delta} + t^*]^2} t \left(1 + \frac{2c}{1 + t^*} \right)^2 \\ &< \frac{2Dv(n-1)^2 a_{\max}^2 \alpha^2 t}{[(t+1)^{1-\delta} + t^*]^2} \left(1 + \frac{c}{\alpha(n-1)a_{\max}} \right)^2. \end{aligned} \quad (4.26)$$

Next we consider the value of $\sum_{i=1}^t \|Z(i)\|$. Using (3.6) and (4.23) we get

$$\begin{aligned} \|Z(i)\| &= \sqrt{V[a(i)\Phi(t, i+1)E\hat{w}(i)]} \\ &< a(i)\sqrt{V[E\hat{w}(i)]} \frac{2c + i^{1-\delta} + t^*}{(t+1)^{1-\delta} + t^*}. \end{aligned} \quad (4.27)$$

By (A5) and (4.25), $V[E\hat{w}(i)] \leq (n-1)a_{\max}^2 \bar{C}^2/i$. From this and (4.27), it can be deduced that

$$\begin{aligned} \sum_{i=1}^t \|Z(i)\| &< \sum_{i=1}^t a(i) \sqrt{\frac{(n-1)a_{\max}^2 \bar{C}^2}{i}} \cdot \frac{2c + i^{1-\delta} + t^*}{(t+1)^{1-\delta} + t^*} \\ &= \frac{\alpha(n-1)a_{\max} \bar{C}}{(t+1)^{1-\delta} + t^*} \sum_{i=1}^t \frac{2c + i^{1-\delta} + t^*}{(i^{1-\delta} + t^*)\sqrt{i}} \\ &< \frac{\alpha(n-1)a_{\max} \bar{C}}{(t+1)^{1-\delta} + t^*} \sum_{i=1}^t \frac{1}{\sqrt{i}} \left(1 + \frac{c}{\alpha(n-1)a_{\max}} \right) \\ &< \frac{2\sqrt{t}\alpha(n-1)a_{\max} \bar{C}}{(t+1)^{1-\delta} + t^*} \left(1 + \frac{c}{\alpha(n-1)a_{\max}} \right). \end{aligned} \quad (4.28)$$

Finally, by (4.23) we have

$$V(\Phi(t, 1)x(1)) < V[x(1)] \left(\frac{2c + t^*}{(t+1)^{1-\delta} + t^*} \right)^2. \quad (4.29)$$

Taking (4.26), (4.28) and (4.29) into (4.21) yields (4.19).

With the similar process of (3.12)-(3.14) we can get the system will reach consensus. \square

5 Main Results

This section will summarize the results of Sections 3 and 4 and give some accurate results for consensus conditions and convergence rate. Also, we will provide an application of our results to our system with stochastic network topologies.

5.1 Critical Conditions for Consensus

The consensus conditions of the first-order average-consensus protocols with deterministic topologies and additive noises have been investigated by some literature. However, the current

best condition of topologies for consensus is the uniform joint-connectivity [31, 38, 40]. On the other hand, if this type protocol contains no noise, it can reach consensus under the infinite joint-connectivity condition[30]. There exists a huge gap between these two consensus conditions. This paper proposes an extensible joint-connectivity condition which is an intermediate condition between the uniform joint-connectivity and infinite joint-connectivity. Under our new condition we investigate a basic problem: what is the critical extensible exponent we can control the system reaching consensus? Note that there does not exist a central controller who knows the global information during our system's evolution, and the consensus control means to design off-line gains $a(t)$ such that all the agents achieve a same state. Recall that \mathcal{G} is the set of the topology sequences satisfying (A1)-(A2). We firstly give an answer under the i.i.d. noises:

Theorem 5.1 *Suppose the noises satisfy (A4). Then we can choose the gain sequence $\{a(t)\}$ a priori such that for any topology sequence $\{\mathcal{G}(t)\} \in \mathcal{G}$ and any initial state, system (2.1)-(2.2) reaches consensus in mean square if and only if $\delta \leq 1/2$, where δ is the extensible exponent appearing in (A1).*

Proof Immediate from Corollary 3.1 and Theorem 3.2. □

Remark 5.1 *Theorem 5.1 says if the network topologies satisfy (A1) and (A2) with $\delta \leq 1/2$, there must exist controls $\{a(t)\}$ such that the system (2.1)-(2.2) reaches consensus in mean square. On the other hand if the topologies satisfy (A1) and (A2) with $\delta > 1/2$, for any $\{a(t)\}$ there must exist some topologies such that the system cannot reach consensus in mean square.*

Remark 5.2 *Though the balanced assumption (A2) needed by Lemma 3.1 had been widely used[22, 28, 31, 37, 39, 40], it may be too strong to the practical applications. In fact, this assumption may not necessary for our results because it has been conjectured that without (A2) Lemma 3.1 still holds with some modifications, see the open problem of the convergence speed of inhomogeneous Markov chain (Problem 1.1 in [53]). If this open problem has been solved, then without (A2) almost all the results of this paper should still hold with slight adjustments.*

We also provide a critical condition for the expectation of the $\tilde{\rho}$ -mixing noises. Let \mathcal{W} be the set of the noise sequences satisfying (A3).

Theorem 5.2 *We can choose the gain sequence $\{a(t)\}$ a priori such that for any topology sequence $\{\mathcal{G}(t)\} \in \mathcal{G}$, any noise sequence $\{w_{ji}(t)\} \in \mathcal{W}$, and any initial state, the system (2.1)-(2.2) will reach consensus in mean square if and only if $\delta \leq 1/2$ and $\varepsilon > 0$, where δ and ε are the constants appearing in (A1) and (A3) respectively.*

Proof Immediate from Theorems 3.1 and 3.3. □

5.2 Estimation of Fastest Convergence Rates

The convergence speed is one of the most important performances of distributed consensus algorithms for networked systems, however most of the existing work focuses on the noise-free algorithms [22–25, 27, 28, 45, 46] where the control gains $\{a(t)\}$ are taken as a constant. Among these research some literature tries to maximize the convergence speed by optimizing the weighted network topology[23, 47]. There also exist few works considering the convergence speed of the distributed consensus algorithms with additive noises. [41, 42, 44]. Nevertheless, it appears that our paper is the first to investigate the fastest convergence rate of this

type protocol with time-varying network topologies and additive noises. Remark that in our system each node only knows its own and neighbors' information and the network topologies cannot be real-time controlled.

We recall that in this paper the fastest convergence rate of consensus at time t is the minimal value of $E[V(x(t))]$ among all the gain functions $a(1) \geq 0, a(2) \geq 0, \dots, a(t-1) \geq 0$ which are the only controllable variables, and $\rho_1(t)$ defined by (4.1) is the fastest convergence rate concerning with the best topologies.

Theorem 5.3 *Suppose (A4) is satisfied. Then for any inconsistent initial state $\rho_1(t) = \Theta\left(\frac{1}{t}\right)$ under protocol (2.1)-(2.2). Also, for all $\{\mathcal{G}(t)\}$ satisfying (A1)-(A2) with $\delta = 0$, choosing $a(t) = \frac{\alpha}{t+t^*}$ can get the convergence rate is $\Theta\left(\frac{1}{t}\right)$, where α and t^* are two constants satisfying (4.20).*

Proof By Theorem 4.2 there exist two constants $t^0 > 0$ and $C_L > 0$ depending on the system parameters and initial states only such that $\rho_1(t) \geq C_L/t$ for all $t \geq t^0$. Also, for all the topologies satisfying (A1)-(A2) with $\delta = 0$, if we choose $a(t) = \frac{\alpha}{t+t^*}$, by Theorem 4.3 and its proof there exist two constants $t^1 > 0$ and $C_U > 0$ depending on the system parameters and initial states only such that $\rho_1(t) \leq C_U/t$ for all $t \geq t^1$. \square

Remark 5.3 *This theorem indicates that for any balanced and uniformly jointly connected topologies, the fastest convergence speed is $\Theta\left(\frac{1}{t}\right)$.*

Remark 5.4 *We can just evaluate the fastest convergence rate to the accurate order. In fact, it is conjectured that $\rho_1(t) = \frac{b_1}{t}(1 + o(1))$ under (A4), where b_1 is a constant depending on n , a_{\max} and v only, however its proof is difficult.*

We recall that $\rho_2(t)$ defined by (4.2) denotes the fastest convergence rate with respect to the worst topologies satisfying (A1)-(A2).

Theorem 5.4 *Suppose the topologies $\{\mathcal{G}(t)\}$ satisfy (A1)-(A2) with $\delta < \frac{1}{2}$, and the noises satisfy (A4). Then for any inconsistent initial state $\rho_2(t) = \Theta\left(\frac{1}{t^{1-2\delta}}\right)$ under protocol (2.1)-(2.2). Also, choosing $a(t) = \frac{\alpha}{t^{1-\delta}+t^*}$ can get the convergence rate is $\Theta\left(\frac{1}{t^{1-2\delta}}\right)$, where α and t^* are two constants satisfying (4.20).*

Proof With the fact of $\rho_2(t)$ and $\rho_3(t)$ are equivalent under (A4), our results can be deduced directly from Theorems 4.1 and 4.3. \square

Similar to Theorem 5.4, by Theorems 4.1 and 4.3 we also have the following theorem:

Theorem 5.5 *Suppose the topologies $\{\mathcal{G}(t)\}$ satisfy (A1)-(A2) with $\delta < \frac{1}{2}$, and the noises satisfy (A5). Then for any inconsistent initial state $\rho_3(t) = \Theta\left(\frac{1}{t^{1-2\delta}}\right)$ under protocol (2.1)-(2.2). Also, choosing $a(t) = \frac{\alpha}{t^{1-\delta}+t^*}$ can get the convergence rate is $\Theta\left(\frac{1}{t^{1-2\delta}}\right)$, where α and t^* are two constants satisfying (4.20).*

Remark 5.5 *Theorems 5.4 and 5.5 establish a relation between the extensible exponent δ and the fastest convergence speed. Similar to Remark 5.4, it is conjectured that $\rho_i(t) = \frac{b_i}{t^{1-2\delta}}(1 + o(1))$, $i = 2, 3$, where b_i is a constant depending on n , a_{\max} , c , δ and the corresponding noises parameters only, however their proofs are difficult.*

5.3 An Application to Stochastic Network Topology

As mentioned in Subsection 3.1, the extensible joint-connectivity can be used to analyze the system with complex stochastic topologies. In this subsection we assume the network topology $\{\mathcal{G}(t)\}_{t \geq 1} = \{(\mathcal{V}, \mathcal{E}(t), \mathcal{A})\}_{t \geq 1}$ is a stochastic process, so the σ -algebra \mathcal{F}_t of the system (2.1)-(2.2) is extended to $\sigma(w_{ji}(k), 1 \leq k \leq t, 1 \leq i \leq n, j \in \mathcal{N}_i(k), \mathcal{G}(k), 1 \leq k \leq t)$. In detail, $\mathcal{G}(t)$ randomly takes one element of \mathcal{B} satisfying:

- (A1') i) All the elements of \mathcal{B} are the balanced graphs;
- ii) $\{\mathcal{G}(t)\}$ is independent of the noises $\{w_{ji}(t)\}$;

Also, we introduce the following two different assumptions for the connectivity of $\{\mathcal{G}(t)\}$, where our result must use one of them.

(A2') There exist three constants $m^* \in \mathbb{Z}^+$, $\mu \in (0, 1/2]$ and $p_{\min} > 0$ such that: For each time $t > m^*$ there exists a subset $\mathcal{B}(t) \subseteq \mathcal{B}$ satisfying i) $\cup_{\mathcal{G} \in \mathcal{B}(t)} \mathcal{G}$ is strongly connected, and ii)

$$P(\mathcal{G}(t) = \mathcal{G} | F) \geq p_{\min} t^{-\frac{\mu}{n-1} \log \frac{1}{n-1} t}, \quad \forall \mathcal{G} \in \mathcal{B}(t), F \in \mathcal{F}_{t-m^*}.$$

(A2*) There exist three constants $m^* \in \mathbb{Z}^+$, $\mu \in (0, 1/2]$ and $p_{\min} > 0$ such that

$$P(\mathcal{G}(t) \text{ is strongly connected} | F) \geq p_{\min} t^{-\mu \log t}, \quad \forall t > m^*, F \in \mathcal{F}_{t-m^*}.$$

Remark 5.6 The assumption (A2') (or (A2*)) includes a wide class of non stationary and strongly correlated random matrix sequence $\{\mathcal{G}(t)\}$, of course including the ergodic and stationary Markov process used in some previous papers[29, 37, 41].

We use the following assumption for the noise as same as (A3) but using u instead of δ :

(A3') The sequence of noises $\{w_{ji}(t), t \geq 1, i = 1, \dots, n, j \in \mathcal{N}_i(t)\}$ are $\tilde{\rho}$ -mixing satisfying: (i) $\sup_{i,j,t} \text{Var}(w_{ji}(t)) < \infty$; (ii) there exists a constant $\varepsilon > 0$ such that

$$\max_{1 \leq i \leq n, j \in \mathcal{N}_i(t)} |E[w_{ji}(t)]| = O(t^{-\mu} \log^{-\varepsilon} t),$$

where μ is the same constant appearing in (A2') (or (A2*)).

From Theorem 3.1 we get the following result for the case of stochastic network topology.

Corollary 5.1 Consider the system given by (2.1)-(2.2) with stochastic network topology satisfying (A1') and (A2') (or (A2*)). Assume the noises satisfy (A3'). Then for any initial state the system will reach consensus in mean square if we choose $a(t) = \frac{\alpha}{t^{1-\mu} \log^\gamma(1+t)}$, where $\alpha > 0$ and $\gamma \in (\max\{\frac{1}{2}, 1 - \varepsilon\}, 1]$ are two arbitrary constants, and μ and ε are the same constants appearing in (A2') (or (A2*)) and (A3').

Proof The main idea of this proof is to show the extensible joint-connectivity can be satisfied with a probability arbitrarily close to 1 for the stochastic network topology. The detailed discussion is put in the appendix. \square

Remark 5.7 Compared to many previous papers which use the assumptions of i.i.d. or stationary Markov process for the stochastic network topologies of the consensus protocols[25–29, 37, 41], this result may be more practical to some applications. For example, for a mobile wireless sensor networks or a multi robot system located in the complicated terrain, the network topologies are a non stationary and strongly correlated random sequence because the

communication probability between two agents depends on the terrain and their current distance, where their current distance depends on not only their distance of the last moment but also their velocities. This kind topology sequence should not be a Markov process but may satisfy our assumptions.

6 Conclusions

Consensus behavior of multi-agent systems has drawn substantial interests over two past decades. However, some key problems remain unsolved, including the fastest convergence speeds and critical consensus conditions of topologies. This paper addresses these problems based on a first-order average-consensus protocol with switching topologies and additive noises. We first propose an extensible joint-connectivity condition for topologies, which can increase the robustness greatly compared to the uniform joint-connectivity condition used widely. Using stochastic approximation methods and under our new condition, we establish some critical consensus conditions for the network topologies and noises, and provide the fastest convergence speeds with respect to the best and worst topologies respectively. Our results give a quantitative description to the relation between the convergence speed and the connectivity of the topologies. Also, as an application we give a consensus analysis for our system with non stationary and strongly correlated stochastic topologies.

References

1. Y. Kuramoto, In *International Symposium on Mathematical Problems in Theoretical Physics*, volume 39 of *Lecture Notes in Physics*, page 420. Springer, New York, 1975.
2. Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*, Springer, Berlin, 1984.
3. J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente *et al.*, “The Kuramoto model: A simple paradigm for synchronization phenomena,” *Rev. Mod. phys.*, vol. 77, no. 1, pp.137-185, 2004.
4. J. Buhl, D. J. T. Sumpter, I. D. Couzin *et al.*, “From disorder to order in marching locusts,” *Science*, vol. 312, no. 5778, pp.1402-1406, 2006.
5. T. Vicsek, A. Czirok, E.B. Jacob *et al.*, “Novel type of phase transition in a system of self-driven particles,” *Phys. Rev. Lett.*, vol. 75, no. 6, pp. 1226-1229, 1995.
6. C. Reynolds, “Flocks, herds, and schools: A distributed behavioral model,” in *Comput. Graph (ACM SIGGRAPH87 Conf. Proc.)*, vol. 21, Jul. 1987, pp. 25-34.
7. G. Deffuant, D. Neau, F. Amblard and G. Weisbuch, “Mixing beliefs among interacting agents,” *Adv. Complex Syst.*, vol. 3, pp. 87-98, 2000.
8. R. Hegselmann and U. Krause, “Opinion dynamics and bounded confidence: Models, analysis and simulation,” *JASSS*, vol. 5, no. 3, 2002.
9. A. Jadbabaie, J. Lin and A.S. Morse, “Coordination of groups of mobile autonomous agents using nearest neighbor rules,” *IEEE Trans. Auto. Control*, vol. 48, no. 9, pp. 988-1001, 2003.

10. F. Cucker and S. Smale, "Emergent Behavior in Flocks," *IEEE Trans. Auto. Control*, vol. 52, no. 5, pp. 852-862, 2007.
11. Z.X. Liu and L. Guo, "Synchronization of multi-agent systems without connectivity assumption," *Automatica*, vol. 45, pp. 2744-2753, 2009.
12. G. Chen, Z.X. Liu and L. Guo, "The Smallest Possible Interaction Radius for Synchronization of Self-Propelled Particles," *SIAM Rev.*, vol. 56, no. 3, pp. 499-521, 2014.
13. A. Rodriguez-Angeles and H. Nijmeijer, "Mutual synchronization of robots via estimated state feedback: a cooperative approach," *IEEE T. Contr. Syst. T.*, vol. 12, no. 4, pp. 542-544, 2004.
14. A. V. Savkin, "Coordinated collective motion of groups of autonomous mobile robots: Analysis of vicsek's model," *IEEE Trans. Autom. Control*, vol. 49, pp. 981-983, June 2004.
15. J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1465-1476, September 2004.
16. W. Ren and R. Beard, *Distributed consensus in multi-vehicle cooperative control: theory and applications*, ser. Communications and Control Engineering Series, Springer-Verlag London, 2008.
17. H. Bai, M. Arcak, and J. T. Wen, "Rigid body attitude coordination without inertial frame information," *Automatica*, vol. 44, no. 12, pp. 3170-3175, 2008.
18. A. Abdessameud and A. Tayebi, "Attitude synchronization of a group of spacecraft without velocity measurements," *IEEE Trans. Autom. Control*, vol. 54, no. 11, pp. 2642-2648, 2009.
19. A. Sarlette, R. Sepulchre, and N. E. Leonard, "Autonomous rigid body attitude synchronization," *Automatica*, vol. 45, no. 2, pp. 572-577, 2009.
20. N. A. Lynch, *Distributed Algorithms*. San Francisco, California: Morgan Kaufmann Publishers, Inc., 1996.
21. R. Olfati-Saber, "Distributed Kalman filter with embedded consensus filters," in *Proc. 44th IEEE Conf. Decision Control Eur. Control Conf.*, Seville, Spain, Dec. 2005, pp. 8179-8184.
22. R. Olfati-Saber and R. M. Murray, "Consensus problem in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520-1533, Sep. 2004.
23. L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, no. 1, pp. 65-78, 2004.
24. A. Olshevsky and J. N. Tsitsiklis, "Convergence speed in distributed consensus and averaging," *SIAM Rev.*, vol. 53, no. 4, pp. 747-772, 2011.
25. Y. Hatano and M. Mesbahi, "Agreement over random networks," *IEEE Trans. Auto. Control*, vol. 50, no. 11, pp. 1867-1872, November 2005.

26. M. Porfiri and D. J. Stilwell, "Consensus seeking over random weighted directed graphs," *IEEE Trans. Auto. Control*, vol. 52, no. 9, pp. 1767-1773, 2007.
27. M. Akar and R. Shorten, "Distributed probabilistic synchronization algorithms for communication networks," *IEEE Trans. Auto. Control*, vol. 53, no. 1, pp. 389-393, 2008.
28. J. Zhou and Q. Wang, "Convergence speed in distributed consensus over dynamically switching random networks," *Automatica*, vol. 45, no. 6, pp. 1455-1461, 2009.
29. A. Tahbaz-Salehi and A. Jadbabaie, "Consensus over ergodic stationary graph processes," *IEEE Trans. Auto. Control*, vol. 55, no. 1, pp. 225-230, 2010.
30. B. Touri and A. Nedić, "On ergodicity, infinite flow, and consensus in random models," *IEEE Trans. Auto. Control*, vol. 56, no. 7, pp. 1593-1605, 2011.
31. B. Touri and A. Nedić, "Distributed consensus over network with noisy links," In *Proc. 12th Int. Conf. Info. Fusion*, Seattle, WA, pp. 146-154, July 2009.
32. S. Patterson, B. Bamieh, and A. El Abbadi, "Convergence rates of distributed average consensus with stochastic link failures," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 880-892, April, 2010.
33. T. C. Aysal and K. E. Barner, "Convergence of consensus models with stochastic disturbances," *IEEE Trans. Info. Theory*, vol. 56, no. 8, pp. 4101-4113, Aug. 2010.
34. G. F. Young, L. Scardovi, and N. E. Leonard, "Robustness of noisy consensus dynamics with directed communication," In *Proc. American Control Conference*, Baltimore, MD, pp. 6312-6317, June 2010.
35. S. S. Stanković, N. Ilić, M. S. Stanković, and K. H. Johansson, "Distributed change detection based on a consensus algorithm," *IEEE Trans. Singal Proces.*, vol. 59, no. 12, pp. 5686-5697, 2011.
36. M. Huang and J.H. Manton, "Coordination and consensus of networked agents with noisy measurements: Stochastic algorithms and asymptotic behavior," *SIAM J. Control Optim.*, vol. 48, no. 1, pp. 134-161, 2009.
37. M. Huang, S. Dey, G. N. Nair, and J. H. Manton, "Stochastic consensus over noisy networks with Markovian and arbitrary switches," *Automatica*, vol. 46, no. 10, pp. 1571-1583, 2010.
38. M. Huang, "Stochastic approximation for consensus: A new approach via ergodic backward products," *IEEE Trans. Auto. Control*, vol. 57, no. 12, pp. 2994-3008, 2012.
39. T. Li and J. F. Zhang, "Mean square average consensus under measurement noises and fixed topologies: Necessary and sufficient conditions," *Automatica*, vol. 45, no. 8, pp. 1929-1936, 2009.
40. T. Li and J. F. Zhang, "Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises," *IEEE Trans. Auto. Control*, vol. 55, no. 9, pp. 2043-2057, 2010.

41. G. Yin, Y. Sun and L.Y. Wang, "Asymptotic properties of consensus-type algorithms for networked systems with regime-switching topologies," *Automatica*, vol. 47, no. 7, pp.1366-1378, 2011.
42. G. Yin, L.Y. Wang, Y. Sun, D. Casbeer, R. Holsapple, and D. Kingston, "Asymptotic optimality for consensus-type stochastic approximation algorithms using iterate averaging," *J. Control Theory A.*, vol. 11, no. 1, pp. 1-9, 2013.
43. L. Y. Wang, A. Syed, G. Yin, A. Pandya and H.W. Zhang, "Coordinated vehicle platoon control: weighted and constrained consensus and communication network topologies," In *Proc. 51st IEEE Conf. Decision Control*, Maui, Hawaii, Dec. 2012, pp. 4057-4062.
44. L.Y. Wang and G. Yin, "Weighted and constrained consensus control with performance optimization for robust distributed UAV Deployments with dynamic information networks," in *Recent Advances in Research on Unmanned Aerial Vehicles*, (Editors: F. Fahroo, L.Y. Wang, and G. Yin), Springer, 2013, pp. 181-206.
45. D. Angeli and P.-A. Bliman, "Tight estimates for convergence of some non-stationary consensus algorithms," *Syst. Control Lett.*, vol. 57, no. 12, pp. 996-1004, 2008.
46. D. Angeli and P.-A. Bliman, "Convergence speed of unsteady distributed consensus: Decay estimate along the settling spanning-trees," *SIAM J. Control Optim.*, vol. 48, no. 1, pp. 1-32, 2009.
47. Y. Kim and M. Mesbahi, "On maximizing the second smallest eigenvalue of a state-dependent graph laplacian," *IEEE Trans. Auto. Control*, vol. 51, no. 1, pp. 116-120, 2006.
48. Ge Chen, "Small noise may diversify collective motion," In print, <http://arxiv.org/abs/1412.0556>.
49. Q.Y. Wu and Y.Y. Jiang, "Some strong limit theorems for $\tilde{\rho}$ -mixing sequences of random variables," *Stat. Probabil. Lett.*, Vol. 78, no. 8, pp. 1017-1023, 2008.
50. Q. M. Shao, "Almost sure invariance principles for mixing sequences of random variables," *Stoch. Proc. Appl.*, vol. 48, no. 2, pp. 319-334, 1993.
51. S. Utev and M. Peligrad, "Maximal inequalities and an invariance principle for a class of weakly dependent random variables," *J. Theor. Probab.*, vol. 16, no. 1, pp. 101-115, 2003.
52. L. Guo, *Time-Varying Stochastic Systems: Stability, Estimation and Control*, Jilin Science and Technology Press (in Chinese), 1993.
53. L. Saloff-Coste and J. Zúñiga, "Merging for time inhomogeneous finite Markov chains, Part I: Singular values and stability," *Electron. J. Probab.*, vol. 14, pp. 1456-1494, 2009.

A Appendix

Proof of Lemma 3.3 We will prove this by induction. For $k = 1$, this result holds evidently.

If this result holds for $k = k'$, then

$$\begin{aligned} t_{k'+1} &\leq t_{k'} + ct_{k'}^\delta \leq (ck')^{\frac{1}{1-\delta}} + c(ck')^{\frac{\delta}{1-\delta}} \\ &= (ck')^{\frac{1}{1-\delta}} (1 + (k')^{-1}) \leq (ck')^{\frac{1}{1-\delta}} (1 + (k')^{-1})^{\frac{1}{1-\delta}} \\ &= (c(k' + 1))^{\frac{1}{1-\delta}}. \end{aligned}$$

By induction, the result holds for all $k \geq 1$. □

Proof of Lemma 3.4 We can compute that

$$y > \int_{t=2}^{\infty} \frac{1}{t \log^\gamma(t)} dt > \sum_{k=1}^{\infty} \frac{1}{\log^\gamma(2^{k+1})} \int_{2^k}^{2^{k+1}} \frac{dt}{t} = \sum_{k=1}^{\infty} \frac{\log 2}{(k+1)^\gamma \log^\gamma(2)} = \infty$$

if $\gamma \leq 1$, and

$$\begin{aligned} y &< \frac{1}{2 \log^\gamma(2)} + \int_{t=2}^{\infty} \frac{1}{t \log^\gamma(t)} dt < \frac{1}{2 \log^\gamma(2)} + \sum_{k=1}^{\infty} \frac{1}{\log^\gamma(2^k)} \int_{2^k}^{2^{k+1}} \frac{dt}{t} \\ &= \frac{1}{2 \log^\gamma(2)} + \sum_{k=1}^{\infty} \frac{\log 2}{k^\gamma \log^\gamma(2)} < \infty \end{aligned}$$

if $\gamma > 1$. □

Before the proof of Lemma 3.5, we need to introduce the following lemma:

Lemma A.1 (Theorem 1.2.22 in [52]) Assume that y_n satisfies

$$y_{n+1} = (1 - a_n)y_n + b_n,$$

and $a_n \in [0, 1)$, $\sum_{n=1}^{\infty} a_n = \infty$, then

$$\lim_{n \rightarrow \infty} y_n = \lim_{n \rightarrow \infty} \frac{b_n}{a_n}, \tag{A.1}$$

if the right side of (A.1) exists.

Proof of Lemma 3.5 Firstly we can compute that

$$\begin{aligned} &\sum_{i=1}^t f^2(i) \prod_{j=k^i}^{\tilde{k}^t-1} (1 - c_2 f(t_{j+1})) \\ &\leq \sum_{l=1}^{\tilde{k}^t-2} \sum_{i=t_l}^{t_{l+1}-1} f^2(i) \prod_{j=l+1}^{\tilde{k}^t-1} (1 - c_2 f(t_{j+1})) + \sum_{i=\tilde{k}^t}^{t-1} f^2(i) \\ &< \sum_{l=1}^{\tilde{k}^t-2} (t_{l+1} - t_l) f^2(t_l) \prod_{j=l+1}^{\tilde{k}^t-1} (1 - c_2 f(t_{j+1})) + (t_{\tilde{k}^t+1} - t_{\tilde{k}^t-1}) f^2(t_{\tilde{k}^t-1}). \end{aligned} \tag{A.2}$$

Let

$$y_m := \sum_{l=1}^{m-2} (t_{l+1} - t_l) f^2(t_l) \prod_{j=l+1}^{m-1} (1 - c_2 f(t_{j+1})),$$

then

$$\begin{aligned} y_{m+1} &= (1 - c_2 f(t_{m+1})) y_m \\ &\quad + (t_m - t_{m-1}) f^2(t_{m-1}) (1 - c_2 f(t_{m+1})) := (1 - a_m) y_m + b_m. \end{aligned}$$

According to Lemmas 3.3 and 3.4 we have

$$\begin{aligned} \sum_{m=3}^{\infty} a_m &= c_2 \sum_{m=4}^{\infty} f(t_m) \geq c_2 \sum_{m=4}^{\infty} \frac{1}{(cm)^{\frac{\beta}{1-\delta}} \log^{\gamma}((cm)^{\frac{1}{1-\delta}} + 1)} \\ &\geq c_2 \sum_{m=4}^{\infty} \frac{1}{cm \log((cm)^{\frac{1}{1-\delta}} + 1)} = \infty. \end{aligned}$$

Also, because $t_m \leq t_{m-1} + ct_{m-1}^{\delta} \leq t_{m-1} + c\sqrt{t_{m-1}}$,

$$\frac{b_m}{a_m} \leq \frac{(t_m - t_{m-1}) f^2(t_{m-1})}{c_2 f(t_{m+1})} = O\left(\frac{t_{m-1}^{\delta}}{t_{m-1}^{\beta} \log^{\gamma}(t_{m-1})}\right) \rightarrow 0 \text{ as } m \rightarrow \infty.$$

Thus, by Lemma A.1 we have $\lim_{m \rightarrow \infty} y_m = 0$. Take this and the following equation

$$(t_{m+1} - t_{m-1}) f^2(t_{m-1}) = O\left(\frac{t_{m-1}^{\delta}}{t_{m-1}^{2\beta} \log^{2\gamma}(t_{m-1})}\right) \rightarrow 0 \text{ as } m \rightarrow \infty$$

into (A.2) we can get (3.2).

With the similar discussion (3.3) can be obtained. \square

Proof of Theorem 3.3 If $\delta > 1/2$, by Theorem 3.2 our result is obtained directly, so we just need consider the case of $\delta \leq 1/2$. We assume the noises $\{w_{ji}(t)\}$ are mutually independent and choose t_k, t^* and $\mathcal{G}(t)$ as same as the proof of Theorem 3.2. Let $\hat{w}^*(i) := \hat{w}(i) - E[\hat{w}(i)]$, using (3.18) we can get

$$\begin{aligned} &x(t+1) - x_{\text{ave}}(t+1)\mathbb{1} \\ &= \tilde{\Phi}(t, t^*)x(t^*) + \sum_{i=t^*}^{t-1} a(i)\tilde{\Phi}(t, i+1)\hat{w}^*(i) + a(t)[\hat{w}^*(t) - (\pi\hat{w}^*(t))\mathbb{1}] \\ &\quad + \sum_{i=t^*}^{t-1} a(i)\tilde{\Phi}(t, i+1)E\hat{w}(i) + a(t)[E\hat{w}(t) - (\pi E\hat{w}(t))\mathbb{1}], \end{aligned}$$

then similar to (3.19) we have

$$\begin{aligned} &E[V(x(t+1))] \\ &\geq E\left\|\tilde{\Phi}(t, t^*)x(t^*) + \sum_{i=t^*}^{t-1} a(i)\tilde{\Phi}(t, i+1)E\hat{w}(i) + a(t)[E\hat{w}(t) - (\pi E\hat{w}(t))\mathbb{1}]\right\|^2 \\ &\quad + \sum_{i=t^*}^{t-1} a^2(i)E\|\tilde{\Phi}(t, i+1)\hat{w}^*(i)\|^2. \end{aligned} \tag{A.3}$$

To reach consensus all the items of the right side of (A.3) must converge to 0. We will show this is impossible by contradiction.

First, by (3.21),

$$E\|\tilde{\Phi}(t, i+1)\hat{w}^*(i)\|^2 \geq \left(\frac{1}{2} - \frac{1}{n}\right)^2 (b_{i+1}^t)^2 \text{Var}(w_{21}(i)).$$

so similar to (3.22) we have $\lim_{t \rightarrow \infty} b_{t^*}^t = 0$. Combining this with (3.17) yields $\tilde{\Phi}(t, t^*)x(t^*) \rightarrow 0$. Also, to reach consensus, $\lim_{t \rightarrow \infty} a(t) = 0$. Thus, by (A.3), to reach consensus

$$y(t) := \sum_{i=t^*}^t a(i)\tilde{\Phi}(t+1, i+1)E\hat{w}(i) \rightarrow 0 \text{ as } t \rightarrow \infty. \quad (\text{A.4})$$

For $t > 1$, we choose $Ew_{21}(t) = \frac{1}{t^\delta \log^\varepsilon(t)}$, and $Ew_{ji}(t) = 0$ for $(j, i) \neq (2, 1)$. We see this choice satisfies the assumption (A3). With (3.20) we have

$$\begin{aligned} y_1(t) &= \sum_{i=t^*}^t a(i)b_{i+1}^{t+1} \left[\frac{E\hat{w}_1(i) - E\hat{w}_2(i)}{2} c_{i+1}^{t+1} + \frac{E\hat{w}_1(i) + E\hat{w}_1(i)}{2} - \pi\hat{w}(i) \right] \\ &= \sum_{i=t^*}^t a(i)b_{i+1}^{t+1} \left[\frac{c_{i+1}^{t+1}}{2i^\delta \log^\varepsilon(i)} + \frac{1}{2i^\delta \log^\varepsilon(i)} - \frac{1}{ni^\delta \log^\varepsilon(i)} \right] \\ &\geq \left(\frac{1}{2} - \frac{1}{n}\right) \sum_{i=t^*}^t \frac{a(i)b_{i+1}^{t+1}}{i^\delta \log^\varepsilon(i)}. \end{aligned} \quad (\text{A.5})$$

With similar process of (3.23) we can get

$$\begin{aligned} \sum_{i=t^*}^t \frac{a(i)b_{i+1}^{t+1}}{i^\delta \log^\varepsilon(i)} &> \frac{1}{4} \sum_{k=k'}^{\tilde{k}^t-1} \frac{\lfloor ct_k^\delta \rfloor a(t_k^*)}{t_{k+1}^\delta \log^\varepsilon(t_{k+1})} \prod_{j=k+1}^{\tilde{k}^t-1} [1 - na(t_j^*)] \\ &\geq \frac{1}{4} \sum_{k=k'}^{\tilde{k}^t-1} \frac{(ct_k^\delta - 1)a(t_k^*)}{(t_k + ct_k^\delta)^\delta} \prod_{j=k+1}^{\tilde{k}^t-1} [1 - na(t_j^*)] \\ &\geq \frac{(ct_{k'}^\delta - 1)}{4(t_{k'} + ct_{k'}^\delta)^\delta} \sum_{k=k'}^{\tilde{k}^t-1} a(t_k^*) \prod_{j=k+1}^{\tilde{k}^t-1} [1 - na(t_j^*)] \\ &= \frac{(ct_{k'}^\delta - 1)}{4n(t_{k'} + ct_{k'}^\delta)^\delta} \left(1 - \prod_{j=k'}^{\tilde{k}^t-1} [1 - na(t_j^*)] \right) \\ &\rightarrow \frac{(ct_{k'}^\delta - 1)}{4n(t_{k'} + ct_{k'}^\delta)^\delta} \text{ as } t \rightarrow \infty, \end{aligned} \quad (\text{A.6})$$

where the last line uses the fact of

$$\prod_{j=k'}^{\infty} [1 - na(t_j^*)] = \lim_{t \rightarrow \infty} b_{t^*}^t = 0.$$

substituting (A.6) into (A.5), we conclude that the system cannot reach consensus in mean square. \square

Proof of Lemma 4.1 Take $\tilde{y} = y - y_{\text{ave}}\mathbb{1}$, $\tilde{x} = x - x_{\text{ave}}\mathbb{1}$ and $\pi = \frac{1}{n}\mathbb{1}'$. Then

$$\begin{aligned}\tilde{y} &= y - (\pi y)\mathbb{1} = (I - aL)x - [\pi(I - aL)x]\mathbb{1} \\ &= x - (\pi x)\mathbb{1} - a[Lx - (\pi Lx)\mathbb{1}] = \tilde{x} - a[L\tilde{x} - (\pi L\tilde{x})\mathbb{1}],\end{aligned}$$

where the last inequality uses the fact of $L\mathbb{1} = 0$. Combining this with $\tilde{x}'\mathbb{1} = 0$ we can get

$$\begin{aligned}V(y) &= \|\tilde{y}\|^2 = \tilde{y}'\tilde{y} \geq \|\tilde{x}\|^2 - a\tilde{x}'[L\tilde{x} - (\pi L\tilde{x})\mathbb{1}] - a[L\tilde{x} - (\pi L\tilde{x})\mathbb{1}]'\tilde{x} \\ &= \|\tilde{x}\|^2 - a\tilde{x}'(L + L')\tilde{x} \geq \|\tilde{x}\|^2 - a\lambda_{\max}(L + L')\|\tilde{x}\|^2 \\ &= (1 - a\lambda_{\max}(L + L'))V(x).\end{aligned}$$

□

Proof of Lemma 4.2 First, we will show that

$$t_{k^i+j} \leq \left(c(j+1) + i^{1-\delta}\right)^{1/(1-\delta)} \quad (\text{A.7})$$

for $j \geq -1$. Similar to Lemma 3.3, we will prove this by induction.

By the definition of k^i we have $t_{k^i-1} \leq i$. So (A.7) holds for $j = -1$. Also, if (A.7) holds for $j \geq -1$, then

$$\begin{aligned}t_{k^i+j+1} &\leq t_{k^i+j} + ct_{k^i+j}^\delta \\ &\leq \left(c(j+1) + i^{1-\delta}\right)^{\frac{1}{1-\delta}} + c\left(c(j+1) + i^{1-\delta}\right)^{\frac{\delta}{1-\delta}} \\ &= c^{\frac{1}{1-\delta}}\left(j+1 + \frac{i^{1-\delta}}{c}\right)^{\frac{1}{1-\delta}}\left[1 + \left(j+1 + \frac{i^{1-\delta}}{c}\right)^{-1}\right] \\ &\leq c^{\frac{1}{1-\delta}}\left(j+1 + \frac{i^{1-\delta}}{c}\right)^{\frac{1}{1-\delta}}\left[1 + \left(j+1 + \frac{i^{1-\delta}}{c}\right)^{-1}\right]^{\frac{1}{1-\delta}} \\ &= \left(c(j+2) + i^{1-\delta}\right)^{\frac{1}{1-\delta}}.\end{aligned}$$

Hence, (A.7) also holds for $j+1$. By induction we have (A.7) holds for all $j \geq -1$.

Thus, using (A.7) we can get

$$t_{k^i + \lfloor \frac{(t+1)^{1-\delta} - i^{1-\delta}}{c} \rfloor - 1} \leq \left((t+1)^{1-\delta}\right)^{1/(1-\delta)} = t+1.$$

From the definition of \tilde{t} ,

$$k^i + \lfloor \frac{(t+1)^{1-\delta} - i^{1-\delta}}{c} \rfloor - 1 \leq \tilde{k}^t.$$

This, together with (A.7) and the fact that $\log(1-x) < -x/2$ for $x \in (0, 1)$, implies

$$\begin{aligned}\prod_{j=k^i}^{\tilde{k}^t-1} \left(1 - \frac{c_1}{t_{j+1}^{1-\delta} + t^*}\right) &\leq \prod_{j=0}^{\lfloor \frac{(t+1)^{1-\delta} - i^{1-\delta}}{c} - 2 \rfloor} \left(1 - \frac{c_1}{t_{k^i+j+1}^{1-\delta} + t^*}\right) \\ &\leq \prod_{j=0}^{\lfloor \frac{(t+1)^{1-\delta} - i^{1-\delta}}{c} - 2 \rfloor} \left(1 - \frac{c_1}{c(j+2) + i^{1-\delta} + t^*}\right) \\ &< \exp \left(\sum_{j=0}^{\lfloor \frac{(t+1)^{1-\delta} - i^{1-\delta}}{c} - 2 \rfloor} \frac{-c_1}{2[c(j+2) + i^{1-\delta} + t^*]} \right).\end{aligned} \quad (\text{A.8})$$

Because for any $a > 0$ and integer $b > 0$,

$$\sum_{k=0}^b (k+a)^{-1} > \int_0^{b+1} (x+a)^{-1} dx = \log(b+1+a) - \log a.$$

Hence, according to (A.8) we have

$$\begin{aligned} & \prod_{j=k^i}^{\tilde{k}^{t-1}} \left(1 - \frac{c_1}{t_{j+1}^{1-\delta} + t^*} \right) \\ & < \exp \left\{ \frac{-c_1}{2c} \left[\log \frac{(t+1)^{1-\delta} + t^*}{c} - \log \left(2 + \frac{i^{1-\delta} + t^*}{c} \right) \right] \right\} \\ & = \left(\frac{i^{1-\delta} + 2c + t^*}{(t+1)^{1-\delta} + t^*} \right)^{\frac{c_1}{2c}}. \end{aligned}$$

□

Proof of Corollary 5.1 We firstly assume (A1'), (A2') and (A3') are satisfied. For any $t \geq 1$, by (A2') we can find $\mathcal{T}_i \in \mathcal{B}(t + im^*)$, $1 \leq i \leq n-1$, such that $\cup_{i=1}^{n-1} \mathcal{T}_i$ is strongly connected and

$$\begin{aligned} P(\mathcal{G}(t + im^*) = \mathcal{T}_i | F) \\ \geq p_{\min} \cdot (t + im^*)^{\frac{-\mu}{n-1}} \log^{\frac{1}{n-1}}(t + im^*), \quad \forall 1 \leq i \leq n-1, F \in \mathcal{F}_t, \end{aligned}$$

where m^* , p_{\min} and μ are the same constants appearing in (A2'). From these we have for any $F \in \mathcal{F}_t$,

$$\begin{aligned} & P \left(\bigcup_{k=t+1}^{t+(n-1)m^*} \mathcal{G}(k) \text{ is strongly connected} | F \right) \\ & \geq P \left(\bigcap_{i=1}^{n-1} \{ \mathcal{G}(t + im^*) = \mathcal{T}_i \} | F \right) \\ & = P(\mathcal{G}(t + m^*) = \mathcal{T}_1 | F) \prod_{i=2}^{n-1} P \left(\mathcal{G}(t + im^*) = \mathcal{T}_i | F, \bigcap_{j=1}^{i-1} \{ \mathcal{G}(t + jm^*) = \mathcal{T}_j \} \right) \quad (\text{A.9}) \\ & \geq p_{\min}^{n-1} \prod_{i=1}^{n-1} (t + im^*)^{\frac{-\mu}{n-1}} \log^{\frac{1}{n-1}}(t + im^*). \end{aligned}$$

Let c be a constant large enough. Set $t_k = 1$ and $t_{k+1} = t_k + \lfloor ct_k^\mu \rfloor$. Let E_{t_1, t_2} be the event

of $\bigcup_{k=t_1}^{t_2-1} \mathcal{G}(k)$ is strongly connected. Let $T := (n-1)m^*$. Then for any $F_k \in \mathcal{F}_{t_k-1}$,

$$\begin{aligned}
P(E_{t_k, t_{k+1}} | F_k) &\geq P\left(\bigcup_{i=1}^{\lfloor \frac{t_{k+1}-t_k}{T} \rfloor} E_{t_k+(i-1)T, t_k+iT-1} | F_k\right) \\
&= 1 - P\left(\bigcap_{i=1}^{\lfloor \frac{t_{k+1}-t_k}{T} \rfloor} E_{t_k+(i-1)T, t_k+iT-1}^c | F_k\right) \\
&= 1 - P(E_{t_k, t_k+T-1}^c | F_k) \\
&\quad \cdot \prod_{i=2}^{\lfloor \frac{t_{k+1}-t_k}{T} \rfloor} P\left(E_{t_k+(i-1)T, t_k+iT-1}^c | F_k, \bigcap_{j=1}^{i-1} \{E_{t_k+(j-1)T, t_k+jT-1}^c\}\right) \\
&\geq 1 - \prod_{i=1}^{\lfloor \frac{t_{k+1}-t_k}{T} \rfloor} \left(1 - p_{\min}^{n-1} \prod_{i=1}^{n-1} (t_k + im^*)^{\frac{-\mu}{n-1}} \log^{\frac{1}{n-1}}(t_k + im^*)\right) \\
&> 1 - \left(1 - p_{\min}^{n-1} t_{k+1}^{-\mu} \log t_{k+1}\right)^{\lfloor \frac{t_{k+1}-t_k}{T} \rfloor},
\end{aligned} \tag{A.10}$$

where the second inequality uses (A.9). Let $p_t = p_{\min}^{n-1} t^{-\mu} \log t$. Because $t_{k+1} - t_k = \lfloor ct_k^\mu \rfloor$, from (A.10) we can get

$$\begin{aligned}
P\left(\bigcap_{k=1}^{\infty} E_{t_k, t_{k+1}}\right) &= P(E_{t_1, t_2}) \prod_{k=2}^{\infty} P\left(E_{t_k, t_{k+1}} | \bigcap_{j=1}^{k-1} E_{t_j, t_{j+1}}\right) \\
&> \prod_{k=1}^{\infty} \left(1 - (1 - p_{t_{k+1}})^{\lfloor \frac{t_{k+1}-t_k}{T} \rfloor}\right) = \prod_{k=1}^{\infty} \left(1 - (1 - p_{t_{k+1}})^{\lfloor \frac{\lfloor ct_k^\mu \rfloor}{T} \rfloor}\right) := f(c).
\end{aligned} \tag{A.11}$$

Let us discuss the value of $f(c)$. First we consider

$$\begin{aligned}
g(c) &:= \sum_{k=1}^{\infty} (1 - p_{t_{k+1}})^{\lfloor \frac{\lfloor ct_k^\mu \rfloor}{T} \rfloor} = \sum_{k=1}^{\infty} \exp\left(\left\lfloor \frac{\lfloor ct_k^\mu \rfloor}{T} \right\rfloor \log(1 - p_{t_{k+1}})\right) \\
&< \sum_{k=1}^{\infty} \exp\left(-\left\lfloor \frac{\lfloor ct_k^\mu \rfloor}{T} \right\rfloor p_{t_{k+1}}\right).
\end{aligned} \tag{A.12}$$

By induction we can obtain $t_k > \frac{c}{2} k^{\frac{1}{1-\mu}}$ for all $k \geq 2$, so for large k we have

$$\begin{aligned}
\exp\left(-\left\lfloor \frac{\lfloor ct_k^\mu \rfloor}{T} \right\rfloor p_{t_{k+1}}\right) &= \exp\left(-\left\lfloor \frac{\lfloor ct_k^\mu \rfloor}{T} \right\rfloor (\underline{c} p_{\min})^{n-1} \frac{\log(t_k + \lfloor ct_k^\mu \rfloor)}{(t_k + \lfloor ct_k^\mu \rfloor)^u}\right) \\
&< \exp\left(-\frac{c}{2T} (\underline{c} p_{\min})^{n-1} \log t_k\right) < \left(\frac{c}{2}\right)^{\frac{-c}{2T} (\underline{c} p_{\min})^{n-1}} k^{\frac{-c}{2T(1-\mu)} (\underline{c} p_{\min})^{n-1}}.
\end{aligned}$$

Taking this into (A.12) we can get $\lim_{c \rightarrow \infty} g(c) = 0$. Thus,

$$\lim_{c \rightarrow \infty} f(c) = \lim_{c \rightarrow \infty} e^{-g(c)} = 1. \tag{A.13}$$

One the other hand, by the total probability theorem and (A.11),

$$\begin{aligned}
E[V(x(t))] &= P\left(\bigcap_{k=1}^{\infty} E_{t_k, t_{k+1}}\right) E\left[V(x(t)) \mid \bigcap_{k=1}^{\infty} E_{t_k, t_{k+1}}\right] \\
&+ \left(1 - P\left(\bigcap_{k=1}^{\infty} E_{t_k, t_{k+1}}\right)\right) E\left[V(x(t)) \mid \bigcup_{k=1}^{\infty} E_{t_k, t_{k+1}}^c\right] \\
&\leq E\left[V(x(t)) \mid \bigcap_{k=1}^{\infty} E_{t_k, t_{k+1}}\right] + (1 - f(c)) E\left[V(x(t)) \mid \bigcup_{k=1}^{\infty} E_{t_k, t_{k+1}}^c\right]
\end{aligned} \tag{A.14}$$

for all large c . Because $\bigcap_{k=1}^{\infty} E_{t_k, t_{k+1}}$ implies (A1) is satisfied, by Theorem 3.1 we can get

$$\lim_{t \rightarrow \infty} E\left[V(x(t)) \mid \bigcap_{k=1}^{\infty} E_{t_k, t_{k+1}}\right] = 0 \tag{A.15}$$

for any positive constant c . Furthermore, let F' be the event of $\bigcup_{k=1}^{\infty} E_{t_k, t_{k+1}}^c$, and set $\hat{w}^*(i) := \hat{w}(i) - E[\hat{w}(i)]$. Because

$$V(x(t)) = \sum_{i=1}^n x_i^2(t) - \frac{1}{n} \left(\sum_{i=1}^n x_i(t) \right)^2 \leq \|x(t)\|^2,$$

from (3.4) we have

$$\begin{aligned}
E[V(x(t)) | F'] &\leq E(\|x(t)\|^2 | F') \\
&= E\left[\left\|\Phi(t-1, 1)x(1) + \sum_{i=1}^{t-1} a(i)\Phi(t-1, i+1)E\hat{w}(i)\right\|^2 | F'\right] \\
&\quad + E\left[\left\|\sum_{i=1}^{t-1} a(i)\Phi(t-1, i+1)\hat{w}^*(i)\right\|^2 | F'\right] \\
&\leq n \left(\max_{1 \leq j \leq n} |x_j(1)| + \sum_{i=1}^{t-1} a(i) \max_{1 \leq j \leq n} |E\hat{w}_j(i)| \right)^2 + 2Dn \sum_{i=1}^{t-1} a^2(i) \max_{1 \leq j \leq n} \text{Var}(\hat{w}_j^*(i)) \\
&= O\left(\sum_{i=1}^{t-1} \frac{1}{i \log^{\gamma+\varepsilon}(1+i)}\right)^2 + O\left(\sum_{i=1}^{t-1} \frac{1}{i^{2(1-\mu)} \log^{2\gamma}(1+i)}\right) < \infty \text{ as } t \rightarrow \infty,
\end{aligned}$$

where the second inequality uses Lemma 3.2 and D is the same constant appearing in it, and the last line uses (A3') and Lemma 3.4. We remark that the upper bound of $E[V(x(t)) | F']$ does not depend on the constant c , so by (A.13) we have

$$\lim_{c \rightarrow \infty} \lim_{t \rightarrow \infty} (1 - f(c)) E[V(x(t)) | F'] = 0.$$

Taking this and (A.15) into (A.14) we have $\lim_{t \rightarrow \infty} E[V(x(t))] = 0$ when we let the constant c in the right side of (A.14) grows to infinite.

Finally, let x^* be the same value defined by (3.12). As same as (3.13) and (3.14) we can obtain $|E[x^*]| < \infty$ and $\text{Var}(x^*) < \infty$ respectively.

For the case of the assumptions (A1'), (A2*) and (A3') are satisfied, we can get our result with the same discussion from (A.10) to the previous paragraph but using $T := m^*$ instead of $T := (n-1)m^*$ and p_{\min} instead of p_{\min}^{n-1} . \square